

STIC-Biotech/ChemLib

89421

From: Li, Ruixiang  
Sent: Wednesday, March 19, 2003 10:56 AM  
To: STIC-Biotech/ChemLib  
Subject: Sequence search of Application NO: 09/727,739

Please do a standard search on:

- (i) SEQ ID NOS: 15 and 18 against the interference amino acid databases;
- (ii) SEQ ID NOS: 17 and 19 against both commercial and interference amino acid databases.

Thank you very much!

Ruixiang Li  
GAU 1646  
CM1 10E18  
Mail Box 10D19  
306-0282

RECEIVED  
MAR 19 2003  
BIOCH/CHEM. DIVISION  
(STIC)

Point of Contact  
P. Sheppard

Telephone number: (703) 308-4499

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: \_\_\_\_\_  
Date Completed: 3/25/03  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:

NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)

STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_



GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2003, 11:53:33 ; Search time 21.4378 Seconds  
(without alignments)  
276.816 Million cell updates/sec

Title: US-09-727-739B-15  
Perfect score: 597  
Sequence: 1 MRVSIQHICALALLGLAIC.....PPRRKAGCKNFYWKGTSC 111

Scoring table: BLOSUM62  
Gapop 10.0 ; Gapext 0.5

Searched: 221153 seqs, 53462247 residues

Total number of hits satisfying chosen parameters: 221153

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	160.5	26.9	110	10	US-09-766-396-3
2	160.5	26.9	110	12	US-10-062-375-3
3	105.5	17.7	140	10	US-09-280-030-64
4	88	14.7	200	9	US-10-101-487-53
5	87	14.6	112	10	US-09-766-396-2
6	87	14.6	112	12	US-10-062-375-2
7	83	13.9	85	10	US-09-766-396-6
8	83	13.9	85	12	US-10-062-375-6
9	80.5	13.5	105	10	US-09-766-396-26
10	80.5	13.5	105	12	US-10-062-375-26
11	80.5	13.5	155	9	US-10-066-500-132
12	80.5	13.5	155	9	US-10-028-072-380
13	80.5	13.5	155	9	US-10-121-049-380
14	80.5	13.5	155	9	US-10-123-904-380
15	80.5	13.5	155	9	US-10-140-470-380
16	80.5	13.5	155	9	US-10-175-746-380
17	80.5	13.5	155	9	US-10-176-918-380
18	80.5	13.5	155	9	US-10-176-921-380
19	80.5	13.5	155	9	US-10-002-796-132

20	80.5	13.5	155	9	US-10-066-273-132	Sequence 132, App
21	80.5	13.5	155	9	US-10-066-494-132	Sequence 132, App
22	80.5	13.5	155	9	US-10-137-865-380	Sequence 380, App
23	80.5	13.5	155	9	US-10-140-474-380	Sequence 380, App
24	80.5	13.5	155	9	US-10-142-431-380	Sequence 380, App
25	80.5	13.5	155	9	US-10-143-114-380	Sequence 380, App
26	80.5	13.5	155	9	US-10-066-269-132	Sequence 132, App
27	80.5	13.5	155	9	US-10-140-002-380	Sequence 380, App
28	80.5	13.5	155	9	US-10-066-193-132	Sequence 132, App
29	80.5	13.5	155	9	US-10-066-211-132	Sequence 132, App
30	80.5	13.5	155	9	US-10-142-419-380	Sequence 380, App
31	80.5	13.5	155	9	US-10-123-262-380	Sequence 380, App
32	80.5	13.5	155	9	US-10-142-423-380	Sequence 380, App
33	79.5	13.3	109	10	US-09-766-396-5	Sequence 5, Appl
34	79.5	13.3	109	12	US-10-062-375-5	Sequence 5, Appl
35	78	13.1	15	9	US-10-101-487-52	Sequence 52, Appl
36	78	13.1	29	10	US-09-766-396-7	Sequence 7, Appl
37	78	13.1	29	10	US-09-766-396-11	Sequence 11, Appl
38	78	13.1	29	12	US-10-062-375-7	Sequence 7, Appl
39	78	13.1	29	12	US-10-062-375-11	Sequence 11, Appl
40	78	13.1	84	10	US-09-766-396-10	Sequence 10, Appl
41	78	13.1	84	12	US-10-062-375-10	Sequence 10, Appl
42	77	12.9	14	9	US-10-101-487-37	Sequence 37, Appl
43	77	12.9	14	10	US-09-316-505-2	Sequence 2, Appl
44	67.5	11.3	95	10	US-09-864-761-41361	Sequence 41361, A
45	67	11.2	133	9	US-10-114-893-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1

US-09-766-396-3  
; Sequence 3, Application US/09766396  
; Patent No. US20020013456A1  
; GENERAL INFORMATION:  
; APPLICANT: Sutcliffe, Gregor J.  
; de Lecea, Luis  
; Siggins, George R.  
; Henriksen, Steven J.  
; TITLE OF INVENTION: CORTISTATIN: NEUROPEPTIDES,  
; COMPOSITIONS AND METHODS  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE  
; STREET: 10666 No. US20020013456A1th Torrey Pines Road, TPC-8  
; CITY: La Jolla  
; STATE: California  
; COUNTRY: US  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/766,396  
; FILING DATE: 18-Jan-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/857,389  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Schmonsees, William  
; REGISTRATION NUMBER: 31,796  
; REFERENCE/DOCKET NUMBER: 22908-0002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 324-7041  
; TELEFAX: (415) 324-0638  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 110 amino acids  
; TYPE: amino acid







SEQ ID NO 53  
LENGTH: 200  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic fusion  
OTHER INFORMATION: protein  
US-101-487-53

Query Match 14.7%; Score 88; DB 9; Length 200;  
Best Local Similarity 48.4%; Pred. No. 0.021;  
Matches 15; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 81 KEDLERSVDNLPRLRERKAGCKNFYWKGTSC 111  
Db 170 EEEEEEEEEEEEEEEACGCKNFFWKTFSC 200

RESULT 5

US-09-766-396-2  
Sequence 2, Application US/09766396  
Patent No. US20020013456A1  
GENERAL INFORMATION:  
APPLICANT: Sutcliffe, Gregor J.  
de Lecea, Luis  
Siggins, George R.  
Henriksen, Steven J.  
TITLE OF INVENTION: CORTISTATIN: NEUROPEPTIDES,  
COMPOSITIONS AND METHODS  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE  
STREET: 10666 No. US20020013456A1th Torrey Pines Road, TPC-8  
CITY: La Jolla  
STATE: California  
COUNTRY: US  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/766,396  
FILING DATE: 18-Jan-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/857,389  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Schmonsees, William  
REGISTRATION NUMBER: 31,796  
REFERENCE/DOCKET NUMBER: 22908-0002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-7041  
TELEFAX: (415) 324-0638  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 112 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-766-396-2

Query Match 14.6%; Score 87; DB 10; Length 112;  
Best Local Similarity 25.2%; Pred. No. 0.014;  
Matches 30; Conservative 9; Mismatches 42; Indels 38; Gaps 4;

QY 20 CSQGAASQPDLDLRSRLQLRALAALPHRSGVSR-----WRTF 59  
Db 4 CSTRGKRPSALSLLLLLLSGIAASALPLESGPTGQDSVQDATGGRRTGLTLFLAWH-- 61

QY 60 YPNCPLRWPRKVKG-----POLKAKEDLERSVDNLPRLRERKAGCKNFYWKGTSC 111  
Db 62 -----EWASQSSSTAEGGTPELSKRQ--ERPPLQPPHRDKKPKCNFFWKTFSSC 111

RESULT 6

US-10-062-375-2  
Sequence 2, Application US/10062375  
Patent No. US20020133000A1  
GENERAL INFORMATION:  
APPLICANT: Sutcliffe, Gregor J.  
de Lecea, Luis  
Siggins, George R.  
Henriksen, Steven J.  
TITLE OF INVENTION: CORTISTATIN: NEUROPEPTIDES,  
COMPOSITIONS AND METHODS  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE  
STREET: 10666 No. US20020133000A1th Torrey Pines Road, TPC-8  
CITY: La Jolla  
STATE: California  
COUNTRY: US  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/062,375  
FILING DATE: 30-Jan-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/857,389  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Schmonsees, William  
REGISTRATION NUMBER: 31,796  
REFERENCE/DOCKET NUMBER: 22908-0002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-7041  
TELEFAX: (415) 324-0638  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 112 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-10-062-375-2

Query Match 14.6%; Score 87; DB 12; Length 112;  
Best Local Similarity 25.2%; Pred. No. 0.014;  
Matches 30; Conservative 9; Mismatches 42; Indels 38; Gaps 4;

QY 20 CSQGAASQPDLDLRSRLQLRALAALPHRSGVSR-----WRTF 59  
Db 4 CSTRGKRPSALSLLLLLLSGIAASALPLESGPTGQDSVQDATGGRRTGLTLFLAWH-- 61

QY 60 YPNCPLRWPRKVKG-----POLKAKEDLERSVDNLPRLRERKAGCKNFYWKGTSC 111  
Db 62 -----EWASQSSSTAEGGTPELSKRQ--ERPPLQPPHRDKKPKCNFFWKTFSSC 111

RESULT 7

US-09-766-396-6  
Sequence 6, Application US/09766396  
Patent No. US20020013456A1  
GENERAL INFORMATION:  
APPLICANT: Sutcliffe, Gregor J.  
de Lecea, Luis  
Siggins, George R.



Henriksen, Steven J.  
TITLE OF INVENTION: CORTISTATIN: NEUROPEPTIDES,  
COMPOSITIONS AND METHODS  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE  
STREET: 10666 No. US20020013456A1th Torrey Pines Road, TPC-8  
CITY: La Jolla  
STATE: California  
COUNTRY: US  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/766,396  
FILING DATE: 18-Jan-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/857,389  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Schmonsees, William  
REGISTRATION NUMBER: 31,796  
REFERENCE/DOCKET NUMBER: 22908-0002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-7041  
TELEFAX: (415) 324-0638  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 85 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: C-terminal  
SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-09-766-396-6  
Query Match 13.9%; Score 83; DB 10; Length 85;  
Best Local Similarity 44.4%; Pred. No. 0.029;  
Matches 16; Conservative 5; Mismatches 13; Indels 2; Gaps 1;  
QY 76 POLKAKEDLERSVDNLPERRKAGCKNFYWKGTSC 111  
Db 51 PELSKRQ--ERPPLQQPPHRDKKPKCNFFWKTFFSC 84  
RESULT 8  
US-10-062-375-6  
Sequence 6, Application US/10062375  
Patent No. US20020133000A1  
GENERAL INFORMATION:  
APPLICANT: Sutcliffe, Gregor J.  
de Lecea, Luis  
Siggins, George R.  
Henriksen, Steven J.  
TITLE OF INVENTION: CORTISTATIN: NEUROPEPTIDES,  
COMPOSITIONS AND METHODS  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE  
STREET: 10666 No. US20020133000A1th Torrey Pines Road, TPC-8  
CITY: La Jolla  
STATE: California  
COUNTRY: US  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/062,375  
FILING DATE: 30-Jan-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/857,389  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Schmonsees, William  
REGISTRATION NUMBER: 31,796  
REFERENCE/DOCKET NUMBER: 22908-0002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-7041  
TELEFAX: (415) 324-0638  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 85 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: C-terminal  
SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-10-062-375-6  
Query Match 13.9%; Score 83; DB 12; Length 85;  
Best Local Similarity 44.4%; Pred. No. 0.029;  
Matches 16; Conservative 5; Mismatches 13; Indels 2; Gaps 1;  
QY 76 POLKAKEDLERSVDNLPERRKAGCKNFYWKGTSC 111  
Db 51 PELSKRQ--ERPPLQQPPHRDKKPKCNFFWKTFFSC 84  
RESULT 9  
US-09-766-396-26  
Sequence 26, Application US/09766396  
Patent No. US20020013456A1  
GENERAL INFORMATION:  
APPLICANT: Sutcliffe, Gregor J.  
de Lecea, Luis  
Siggins, George R.  
Henriksen, Steven J.  
TITLE OF INVENTION: CORTISTATIN: NEUROPEPTIDES,  
COMPOSITIONS AND METHODS  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE  
STREET: 10666 No. US20020013456A1th Torrey Pines Road, TPC-8  
CITY: La Jolla  
STATE: California  
COUNTRY: US  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/766,396  
FILING DATE: 18-Jan-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/857,389  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Schmonsees, William  
REGISTRATION NUMBER: 31,796  
REFERENCE/DOCKET NUMBER: 22908-0002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-7041  
TELEFAX: (415) 324-0638  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:



LENGTH: 105 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 26:  
US-09-766-396-26

Query Match 13.5%; Score 80.5; DB 10; Length 105;  
Best Local Similarity 27.7%; Pred. No. 0.07;  
Matches 26; Conservative 12; Mismatches 37; Indels 19; Gaps 4;

QY 37 LLQRALA-AALPHRSGVSRWRTFYPCPLR-----WRPRKVKGPQL-KAK 81  
Db 11 LLSGATATAALPLEGGPTGRDSEHMQEAGIRKSSLLTFLAWWFETWSQASAGPLIGEEA 70

QY 82 EDLERSVDNLPP-----RERKAGCKNFYKGFSTSC 111  
Db 71 REVARROEGAPPOQSARRDRMPCRNFFWKTFSSC 104

RESULT 10  
US-10-062-375-26  
; Sequence 26, Application US/10062375  
; Patent No. US20020133000A1  
; GENERAL INFORMATION:  
; APPLICANT: Sutcliffe, Gregor J.  
; de Lecea, Luis  
; Sliggins, George R.  
; Henriksen, Steven J.  
; TITLE OF INVENTION: CORTISTATIN: NEUROPEPTIDES,  
; COMPOSITIONS AND METHODS  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE  
; STREET: 10666 No. US20020133000A1th Torrey Pines Road, TPC-8  
; CITY: La Jolla  
; STATE: California  
; COUNTRY: US  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/062,375  
; FILING DATE: 30-Jan-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/857,389  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Schmonsees, William  
; REGISTRATION NUMBER: 31,796  
; REFERENCE/DOCKET NUMBER: 22908-0002  
; TELEPHONE: (415) 324-7041  
; TELEFAX: (415) 324-0638  
; INFORMATION FOR SEQ ID NO: 26:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 105 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
; SEQUENCE DESCRIPTION: SEQ ID NO: 26:

US-10-062-375-26

Query Match 13.5%; Score 80.5; DB 12; Length 105;  
Best Local Similarity 27.7%; Pred. No. 0.07;  
Matches 26; Conservative 12; Mismatches 37; Indels 19; Gaps 4;

QY 37 LLQRALA-AALPHRSGVSRWRTFYPCPLR-----WRPRKVKGPQL-KAK 81  
Db 11 LLSGATATAALPLEGGPTGRDSEHMQEAGIRKSSLLTFLAWWFETWSQASAGPLIGEEA 70

QY 82 EDLERSVDNLPP-----RERKAGCKNFYKGFSTSC 111  
Db 71 REVARROEGAPPOQSARRDRMPCRNFFWKTFSSC 104

RESULT 11  
US-10-066-500-132  
; Sequence 132, Application US/10066500  
; Patent No. US20020177165A1  
; GENERAL INFORMATION:  
; APPLICANT: Avi J. Ashkenazi  
; APPLICANT: Kevin P. Baker  
; APPLICANT: David A. Botstein  
; APPLICANT: Luc Desnoyers  
; APPLICANT: Dan L. Eaton  
; APPLICANT: Napoleone Ferrara  
; APPLICANT: Sherman Fong  
; APPLICANT: Wei-Qiang Gao  
; APPLICANT: Hanspeter Gerber  
; APPLICANT: Mary E. Gerritsen  
; APPLICANT: Audrey Goddard  
; APPLICANT: Paul J. Godowski  
; APPLICANT: Austin L. Gurney  
; APPLICANT: Ivar J. Kijavlin  
; APPLICANT: Jennie P. Mather  
; APPLICANT: Mary A. Napier  
; APPLICANT: James Pan  
; APPLICANT: Nicholas F. Paoni  
; APPLICANT: Margaret Ann Roy  
; APPLICANT: Timothy A. Stewart  
; APPLICANT: Daniel Tumas  
; APPLICANT: Colin K. Watanabe  
; APPLICANT: P. Mickey Williams  
; APPLICANT: William I. Wood  
; APPLICANT: Zemin Zang  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3130R1C7  
; CURRENT APPLICATION NUMBER: US/10/066,500  
; FILING DATE: 2002-02-01  
; PRIOR APPLICATION NUMBER: 10/002,796  
; PRIOR FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: 60/056974  
; PRIOR FILING DATE: 1997-08-26  
; PRIOR APPLICATION NUMBER: 60/059115  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059263  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/059588  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/062285  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/062816  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063082  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR APPLICATION NUMBER: 60/063329  
; PRIOR FILING DATE: 1997-10-27  
; PRIOR APPLICATION NUMBER: 60/063733  
; PRIOR FILING DATE: 1997-10-29  
; PRIOR APPLICATION NUMBER: 60/066364  
; PRIOR FILING DATE: 1997-11-21  
; PRIOR APPLICATION NUMBER: 60/066840



;; PRIOR FILING DATE: 1997-11-25  
;; PRIOR APPLICATION NUMBER: 60/069694  
;; PRIOR FILING DATE: 1997-12-16  
;; PRIOR APPLICATION NUMBER: 60/074086  
;; PRIOR FILING DATE: 1998-02-09  
;; PRIOR APPLICATION NUMBER: 60/074092  
;; PRIOR FILING DATE: 1998-02-09  
;; PRIOR APPLICATION NUMBER: 60/079294  
;; PRIOR FILING DATE: 1998-03-25  
;; PRIOR APPLICATION NUMBER: 60/081049  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/095998  
;; PRIOR FILING DATE: 1998-08-10  
;; PRIOR APPLICATION NUMBER: 60/097000  
;; PRIOR FILING DATE: 1998-08-18  
;; PRIOR APPLICATION NUMBER: 60/099601  
;; PRIOR FILING DATE: 1998-09-09  
;; PRIOR APPLICATION NUMBER: 60/099803  
;; PRIOR FILING DATE: 1998-09-10  
;; PRIOR APPLICATION NUMBER: 60/099811  
;; PRIOR FILING DATE: 1998-09-10  
;; PRIOR APPLICATION NUMBER: 60/099812  
;; PRIOR FILING DATE: 1998-09-10  
;; PRIOR APPLICATION NUMBER: 60/100858  
;; PRIOR FILING DATE: 1998-09-17  
;; PRIOR APPLICATION NUMBER: 60/101922  
;; PRIOR FILING DATE: 1998-09-24  
;; PRIOR APPLICATION NUMBER: 60/106032  
;; PRIOR FILING DATE: 1998-10-28  
;; PRIOR APPLICATION NUMBER: 60/109304  
;; PRIOR FILING DATE: 1998-11-20  
;; PRIOR APPLICATION NUMBER: 60/125778  
;; PRIOR FILING DATE: 1999-03-23  
;; PRIOR APPLICATION NUMBER: 60/139695  
;; PRIOR FILING DATE: 1999-06-15  
;; PRIOR APPLICATION NUMBER: 60/145070  
;; PRIOR FILING DATE: 1999-07-20  
;; PRIOR APPLICATION NUMBER: 60/145698  
;; PRIOR FILING DATE: 1999-07-26  
;; PRIOR APPLICATION NUMBER: 60/149396  
;; PRIOR FILING DATE: 1999-08-17  
;; PRIOR APPLICATION NUMBER: 60/169495  
;; PRIOR FILING DATE: 1999-12-07  
;; PRIOR APPLICATION NUMBER: 08/918874  
;; PRIOR FILING DATE: 1997-08-26  
;; PRIOR APPLICATION NUMBER: 08/933821  
;; PRIOR FILING DATE: 1997-09-19  
;; PRIOR APPLICATION NUMBER: 08/960507  
;; PRIOR FILING DATE: 1997-10-29  
;; PRIOR APPLICATION NUMBER: 09/114844  
;; PRIOR FILING DATE: 1998-07-14  
;; PRIOR APPLICATION NUMBER: 09/136801  
;; PRIOR FILING DATE: 1998-08-19  
;; PRIOR APPLICATION NUMBER: 09/136804  
;; PRIOR FILING DATE: 1998-08-19  
;; PRIOR APPLICATION NUMBER: 09/136828  
;; PRIOR FILING DATE: 1998-08-19  
;; PRIOR APPLICATION NUMBER: 09/158342  
;; PRIOR FILING DATE: 1998-09-21  
;; PRIOR APPLICATION NUMBER: 09/180997  
;; PRIOR FILING DATE: 1998-09-10  
;; PRIOR APPLICATION NUMBER: 09/202088  
;; PRIOR FILING DATE: 1998-12-08  
;; PRIOR APPLICATION NUMBER: 09/254311  
;; PRIOR FILING DATE: 1999-03-03  
;; PRIOR APPLICATION NUMBER: 09/254460  
;; PRIOR FILING DATE: 1999-03-09  
;; PRIOR APPLICATION NUMBER: 09/254465  
;; PRIOR FILING DATE: 1999-03-05  
;; PRIOR APPLICATION NUMBER: 09/284663  
;; PRIOR FILING DATE: 1999-04-15  
;; PRIOR APPLICATION NUMBER: 09/332928  
;; PRIOR FILING DATE: 1999-06-14

;; PRIOR APPLICATION NUMBER: 09/332929  
;; PRIOR FILING DATE: 1999-06-14  
;; PRIOR APPLICATION NUMBER: 09/333075  
;; PRIOR FILING DATE: 1999-06-14  
;; PRIOR APPLICATION NUMBER: 09/333077  
;; PRIOR FILING DATE: 1999-06-14  
;; PRIOR APPLICATION NUMBER: 09/380137  
;; PRIOR FILING DATE: 1999-08-25  
;; PRIOR APPLICATION NUMBER: 09/380138  
;; PRIOR FILING DATE: 1999-08-25  
;; PRIOR APPLICATION NUMBER: 09/380139  
;; PRIOR FILING DATE: 1999-08-25  
;; PRIOR APPLICATION NUMBER: 09/403296  
;; PRIOR FILING DATE: 1999-10-18  
;; PRIOR APPLICATION NUMBER: 09/403297  
;; PRIOR FILING DATE: 1999-10-18  
;; PRIOR APPLICATION NUMBER: 09/423741  
;; PRIOR FILING DATE: 1999-11-10  
;; PRIOR APPLICATION NUMBER: 09/423844  
;; PRIOR FILING DATE: 1999-11-12  
;; PRIOR APPLICATION NUMBER: 09/522342  
;; PRIOR FILING DATE: 2000-03-09  
;; PRIOR APPLICATION NUMBER: 09/548815  
;; PRIOR FILING DATE: 2000-04-13  
;; PRIOR APPLICATION NUMBER: 09/664610  
;; PRIOR FILING DATE: 2000-09-18  
;; PRIOR APPLICATION NUMBER: 09/665350  
;; PRIOR FILING DATE: 2000-09-18  
;; PRIOR APPLICATION NUMBER: 09/709238  
;; PRIOR FILING DATE: 2000-11-08  
;; PRIOR APPLICATION NUMBER: 09/767609  
;; PRIOR FILING DATE: 2001-01-22  
;; PRIOR APPLICATION NUMBER: 09/802706  
;; PRIOR FILING DATE: 2001-03-09  
;; PRIOR APPLICATION NUMBER: 09/808689  
;; PRIOR FILING DATE: 2001-03-14  
;; PRIOR APPLICATION NUMBER: 09/866028  
;; PRIOR FILING DATE: 2001-05-25  
;; PRIOR APPLICATION NUMBER: 09/870574  
;; PRIOR FILING DATE: 2001-05-30  
;; PRIOR APPLICATION NUMBER: 09/872035  
;; PRIOR FILING DATE: 2001-06-01  
;; PRIOR APPLICATION NUMBER: 09/886342  
;; PRIOR FILING DATE: 2001-06-19  
;; PRIOR APPLICATION NUMBER: PCT/US98/14552  
;; PRIOR FILING DATE: 1998-07-14  
;; PRIOR APPLICATION NUMBER: PCT/US98/18824  
;; PRIOR FILING DATE: 1998-09-10  
;; PRIOR APPLICATION NUMBER: PCT/US98/19093  
;; PRIOR FILING DATE: 1998-09-14  
;; PRIOR APPLICATION NUMBER: PCT/US98/19330  
;; PRIOR FILING DATE: 1998-09-16  
;; PRIOR APPLICATION NUMBER: PCT/US98/19437  
;; PRIOR FILING DATE: 1998-09-17  
;; PRIOR APPLICATION NUMBER: PCT/US98/24855  
;; PRIOR FILING DATE: 1998-11-20  
;; PRIOR APPLICATION NUMBER: PCT/US98/25108  
;; PRIOR FILING DATE: 1998-12-01  
;; PRIOR APPLICATION NUMBER: PCT/US98/25190  
;; PRIOR FILING DATE: 1998-11-25  
;; PRIOR APPLICATION NUMBER: PCT/US99/05028  
;; PRIOR FILING DATE: 1999-03-08  
;; PRIOR APPLICATION NUMBER: PCT/US99/12252  
;; PRIOR FILING DATE: 1999-06-02  
;; PRIOR APPLICATION NUMBER: PCT/US99/20111  
;; PRIOR FILING DATE: 1999-09-01  
;; PRIOR APPLICATION NUMBER: PCT/US99/20594  
;; PRIOR FILING DATE: 1999-09-08  
;; PRIOR APPLICATION NUMBER: PCT/US99/21090  
;; PRIOR FILING DATE: 1999-09-15  
;; PRIOR APPLICATION NUMBER: PCT/US99/21547

Query Match

13.5%; Score 80.5; DB 9; Length 155;



Best Local Similarity 27.7%; Pred. No. 0.11;  
Matches 26; Conservative 12; Mismatches 37; Indels 19; Gaps 4;  
QY 37 LLQRALA-AALPHRSGVSEWRWTFYPCPCLR-----WRPRKVKGPQL-KAK 81  
DB 61 LLSGATATAALPLEGGPTGRDSEHMQEAAGIRKSILLTFLAWFEWTSQASAGPLIGEEA 120  
QY 82 EDLERSVDNLPP-----RERKAGCKNFYWKFTSC 111  
DB 121 REVARROEGAPPOQSARRDRMPCRNFFWKTFSSC 154

RESULT 12

US-10-028-072-380  
; Sequence 380, Application US/10028072  
; Publication No. US20030004311A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang

TITLE OF INVENTION:

FILE REFERENCE:

CURRENT APPLICATION NUMBER: US/10/028,072

CURRENT FILING DATE: 2001-12-19

PRIOR APPLICATION NUMBER: 60/049911

PRIOR FILING DATE: 1997-06-18

PRIOR APPLICATION NUMBER: 60/056974

PRIOR FILING DATE: 1997-08-26

PRIOR APPLICATION NUMBER: 60/059113

PRIOR FILING DATE: 1997-09-17

PRIOR APPLICATION NUMBER: 60/059115

PRIOR FILING DATE: 1997-09-17

PRIOR APPLICATION NUMBER: 60/059117

PRIOR FILING DATE: 1997-09-17

PRIOR APPLICATION NUMBER: 60/059122

PRIOR FILING DATE: 1997-09-17

PRIOR APPLICATION NUMBER: 60/059184

PRIOR FILING DATE: 1997-09-17

PRIOR APPLICATION NUMBER: 60/059263

PRIOR FILING DATE: 1997-09-18

PRIOR APPLICATION NUMBER: 60/059352

PRIOR FILING DATE: 1997-09-19

PRIOR APPLICATION NUMBER: 60/059588

PRIOR FILING DATE: 1997-09-19

PRIOR APPLICATION NUMBER: 60/059836

PRIOR FILING DATE: 1997-09-24

PRIOR APPLICATION NUMBER: 60/062250

PRIOR FILING DATE: 1997-10-17

PRIOR APPLICATION NUMBER: 60/062285

PRIOR FILING DATE: 1997-10-17

PRIOR APPLICATION NUMBER: 60/062287

PRIOR FILING DATE: 1997-10-17

PRIOR APPLICATION NUMBER: 60/062814

PRIOR FILING DATE: 1997-10-24

PRIOR APPLICATION NUMBER: 60/062816

PRIOR FILING DATE: 1997-10-24

PRIOR APPLICATION NUMBER: 60/063045

PRIOR FILING DATE: 1997-10-24

PRIOR APPLICATION NUMBER: 60/063082

PRIOR FILING DATE: 1997-10-31  
PRIOR APPLICATION NUMBER: 60/063127  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063327  
PRIOR FILING DATE: 1997-10-27  
PRIOR APPLICATION NUMBER: 60/063329  
PRIOR FILING DATE: 1997-10-27  
PRIOR APPLICATION NUMBER: 60/063550  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063561  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063704  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063733  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063735  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063738  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063755  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064248  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/064809  
PRIOR FILING DATE: 1997-11-07  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065846  
PRIOR FILING DATE: 1997-11-17  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/066453  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/066511  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/069212  
PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/069278  
PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/069334  
PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/069694  
PRIOR FILING DATE: 1997-12-16  
PRIOR APPLICATION NUMBER: 60/072320  
PRIOR FILING DATE: 1998-01-23  
PRIOR APPLICATION NUMBER: 60/073612  
PRIOR FILING DATE: 1998-02-04  
PRIOR APPLICATION NUMBER: 60/074086  
PRIOR FILING DATE: 1998-02-09  
PRIOR APPLICATION NUMBER: 60/074092  
PRIOR FILING DATE: 1998-02-09  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-02-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081695  
PRIOR FILING DATE: 1998-04-14  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15



;; PRIOR APPLICATION NUMBER: 60/081818  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/082999  
;; PRIOR FILING DATE: 1998-04-24  
;; PRIOR APPLICATION NUMBER: 60/083322  
;; PRIOR FILING DATE: 1998-04-28  
;; PRIOR APPLICATION NUMBER: 60/083545  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/084600  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084627  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084637  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/085149  
;; PRIOR FILING DATE: 1998-05-12  
;; PRIOR APPLICATION NUMBER: 60/085323  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085338  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085339  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085579  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085697  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085704  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/086414  
;; PRIOR FILING DATE: 1998-05-22  
;; PRIOR APPLICATION NUMBER: 60/086430  
;; PRIOR FILING DATE: 1998-05-22  
;; PRIOR APPLICATION NUMBER: 60/087106  
;; PRIOR FILING DATE: 1998-05-28  
;; PRIOR APPLICATION NUMBER: 60/088026  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088730  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088741  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088810  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088858  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/089532  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089599  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089907  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089947  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/090349  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090429  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090445  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090538  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090863  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07

Query Match 13.5%; Score 80.5; DB 9; Length 155;  
Best Local Similarity 27.7%; Pred. No. 0.11;  
Matches 26; Conservative 12; Mismatches 37; Indels 19; Gaps 4;

QY 37 LLQRALA-AALPHRSVGSERWRTFYPCPLR-----WRPRKVKGPQL-KAK 81  
DB 61 LLSGATAAALPLEGGTGRDSEHMOAAGIRKSSLLTFLAWFWFTSQASAGPLIGEEA 120  
QY 82 EDLERSVDNLPP-----RERKAGCKNFYWKGTSC 111  
DB 121 REVARQEGAPPOQSARRDRMPCRNFFWKTFFSSC 154  
RESULT 13  
US-10-121-049-380  
; Sequence 380, Application US/10121049  
; Publication No. US20030022239A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tamas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3330R1C17  
; CURRENT APPLICATION NUMBER: US/10/121,049  
; CURRENT FILING DATE: 2002-04-12  
; Prior Application removed - See File Wrapper or Palm.  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 380  
; LENGTH: 155  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-121-049-380  
Query Match 13.5%; Score 80.5; DB 9; Length 155;  
Best Local Similarity 27.7%; Pred. No. 0.11;  
Matches 26; Conservative 12; Mismatches 37; Indels 19; Gaps 4;

QY 37 LLQRALA-AALPHRSVGSERWRTFYPCPLR-----WRPRKVKGPQL-KAK 81  
DB 61 LLSGATAAALPLEGGTGRDSEHMOAAGIRKSSLLTFLAWFWFTSQASAGPLIGEEA 120  
QY 82 EDLERSVDNLPP-----RERKAGCKNFYWKGTSC 111  
DB 121 REVARQEGAPPOQSARRDRMPCRNFFWKTFFSSC 154  
RESULT 14  
US-10-123-904-380  
; Sequence 380, Application US/10123904  
; Publication No. US200300222328A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven



APPLICANT: Smith,Victoria  
APPLICANT: Stewart,Timothy A.  
APPLICANT: Tumas,Daniel  
APPLICANT: Watanabe,Colin K  
APPLICANT: Wood,William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3330R1C54  
CURRENT APPLICATION NUMBER: US/10/123,904  
CURRENT FILING DATE: 2002-04-16  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 380  
LENGTH: 155  
TYPE: PRT  
ORGANISM: Homo Sapien  
S-10-123-904-380

Query Match 13.5%; Score 80.5; DB 9; Length 155;  
Best Local Similarity 27.7%; Pred. No. 0.11;  
Matches 26; Conservative 12; Mismatches 37; Indels 19; Gaps 4;

QY 37 LLQRALA-AALPHRSGVSRWTFYNPCCLR-----WRPRKVKGPQL-KAK 81  
Db 61 LLSGATATAALPLEGGPTGRDSEHMQEAGIRKSSLLTFLAWWFETWSQASAGPLIGEEA 120  
QY 82 EDLERSVDNLPP---RERKAGCKNFYWKGTSC 111  
Db 121 REVARROEGAPPOQSARRDRMPCRNFFWKTFSSC 154

RESULT 15  
US-10-140-470-380  
Sequence 380, Application US/10140470  
Publication No. US20030022331a1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3330R1C160  
CURRENT APPLICATION NUMBER: US/10/140,470  
CURRENT FILING DATE: 2002-05-06  
Prior Application removed - See Palm or File Wrapper  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 380  
LENGTH: 155  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-140-470-380

Query Match 13.5%; Score 80.5; DB 9; Length 155;  
Best Local Similarity 27.7%; Pred. No. 0.11;  
Matches 26; Conservative 12; Mismatches 37; Indels 19; Gaps 4;

QY 37 LLQRALA-AALPHRSGVSRWTFYNPCCLR-----WRPRKVKGPQL-KAK 81  
Db 61 LLSGATATAALPLEGGPTGRDSEHMQEAGIRKSSLLTFLAWWFETWSQASAGPLIGEEA 120

Db 61 LLSGATATAALPLEGGPTGRDSEHMQEAGIRKSSLLTFLAWWFETWSQASAGPLIGEEA 120  
QY 82 EDLERSVDNLPP---RERKAGCKNFYWKGTSC 111  
Db 121 REVARROEGAPPOQSARRDRMPCRNFFWKTFSSC 154  
Search completed: March 21, 2003, 11:56:28  
Job time : 23.4378 secs















;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE  
;; STREET: 10666 No. 6074872th Torrey Pines Road, TPC-8  
;; CITY: La Jolla  
;; STATE: California  
;; COUNTRY: US  
;; ZIP: 92037  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/648,322  
;; FILING DATE:  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Fitting, Thomas  
;; REGISTRATION NUMBER: 34,163  
;; REFERENCE/DOCKET NUMBER: 519.0  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (619) 554-2937  
;; TELEFAX: (619) 554-6312  
;; INFORMATION FOR SEQ ID NO: 7:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 29 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; FRAGMENT TYPE: C-terminal  
;; US-08-648-322-7

Query Match 13.1%; Score 78; DB 3; Length 29;  
Best Local Similarity 51.9%; Pred. No. 0.004;  
Matches 14; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 85 ERSVDNLPPEKAGCKNFYWKGTSC 111  
|| || | ||||| |:  
Db 2 ERPLQPPHDKKPKCKNFYWKTFSSC 28

RESULT 8  
US-08-648-322-11  
;; Sequence 11, Application US/08648322  
;; Patent No. 6074872  
;; GENERAL INFORMATION:  
;; APPLICANT: Sutcliffe, Gregor J.  
;; APPLICANT: de Lecea, Luis  
;; TITLE OF INVENTION: CORTISTATIN: NEUROPEPTIDES,  
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS  
;; NUMBER OF SEQUENCES: 24  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE  
;; STREET: 10666 No. 6074872th Torrey Pines Road, TPC-8  
;; CITY: La Jolla  
;; STATE: California  
;; COUNTRY: US  
;; ZIP: 92037  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/648,322  
;; FILING DATE:  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Fitting, Thomas  
;; REGISTRATION NUMBER: 34,163  
;; REFERENCE/DOCKET NUMBER: 519.0  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (619) 554-2937

;; TELEFAX: (619) 554-6312  
;; INFORMATION FOR SEQ ID NO: 11:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 29 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; FRAGMENT TYPE: C-terminal  
;; US-08-648-322-11  
Query Match 13.1%; Score 78; DB 3; Length 29;  
Best Local Similarity 51.9%; Pred. No. 0.004;  
Matches 14; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 85 ERSVDNLPPEKAGCKNFYWKGTSC 111  
|| || | ||||| |:  
Db 2 ERPPQPPHDKKPKCKNFYWKTFSSC 28

RESULT 9  
US-08-648-322-10  
;; Sequence 10, Application US/08648322  
;; Patent No. 6074872  
;; GENERAL INFORMATION:  
;; APPLICANT: Sutcliffe, Gregor J.  
;; APPLICANT: de Lecea, Luis  
;; TITLE OF INVENTION: CORTISTATIN: NEUROPEPTIDES,  
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS  
;; NUMBER OF SEQUENCES: 24  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE  
;; STREET: 10666 No. 6074872th Torrey Pines Road, TPC-8  
;; CITY: La Jolla  
;; STATE: California  
;; COUNTRY: US  
;; ZIP: 92037  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/648,322  
;; FILING DATE:  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Fitting, Thomas  
;; REGISTRATION NUMBER: 34,163  
;; REFERENCE/DOCKET NUMBER: 519.0  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (619) 554-2937  
;; TELEFAX: (619) 554-6312  
;; INFORMATION FOR SEQ ID NO: 10:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 84 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; FRAGMENT TYPE: C-terminal  
;; US-08-648-322-10

Query Match 13.1%; Score 78; DB 3; Length 84;  
Best Local Similarity 51.9%; Pred. No. 0.017;  
Matches 14; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 85 ERSVDNLPPEKAGCKNFYWKGTSC 111  
|| || | ||||| |:  
Db 57 ERPPQPPHDKKPKCKNFYWKTFSSC 83

RESULT 10  
US-08-455-970A-10  
;; Sequence 10, Application US/08455970A







SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-255-272-5

Query Match 12.9%; Score 77; DB 1; Length 14;  
Best Local Similarity 85.7%; Pred. No. 0.002;  
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 98 AGCKNFYWKGTSC 111  
Db 1 AGCKNFFWKFTSC 14

RESULT 13

US-08-416-007-4  
Sequence 4, Application US/08416007  
Patent No. 5693679

GENERAL INFORMATION:  
APPLICANT: Vincent, Jean-Pierre  
APPLICANT: Gaudriault, Georges  
APPLICANT: Beaudet, Alain  
TITLE OF INVENTION: FLUORESCENT SOMATOSTATIN  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804

COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/416,007  
FILING DATE: 04-APR-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 06942/003001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-8906  
TELEX: 200154

INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-416-007-4

Query Match 12.9%; Score 77; DB 1; Length 14;  
Best Local Similarity 85.7%; Pred. No. 0.002;  
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 98 AGCKNFYWKGTSC 111  
Db 1 AGCKNFFWKFTSC 14

RESULT 14

US-08-676-263-11  
Sequence 11, Application US/08676263

Patent No. 5705143  
GENERAL INFORMATION:  
APPLICANT: Bower, Gary R.  
APPLICANT: Forster, Alan M.  
APPLICANT: Riley, Anthony L. M.  
APPLICANT: Storey, Anthony E.  
TITLE OF INVENTION: BIOLOGICAL TARGETING AGENTS  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive/6300 Sears Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/676,263  
FILING DATE: 07-NOV-1996  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 94300224.6  
FILING DATE: 12-JAN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Sharp, Jeffrey S.  
REGISTRATION NUMBER: 31,879  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
TELEX: (312) 474-6600

INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Cross-links  
LOCATION: 3..14  
US-08-676-263-11

Query Match 12.9%; Score 77; DB 1; Length 14;  
Best Local Similarity 85.7%; Pred. No. 0.002;  
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 98 AGCKNFYWKGTSC 111  
Db 1 AGCKNFFWKFTSC 14

RESULT 15

US-08-286-748B-13  
Sequence 13, Application US/08286748B  
Patent No. 5759542

GENERAL INFORMATION:  
APPLICANT: Victor Gurewich  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY  
OF DRUGS BY PLATELETS FOR THE TREATMENT OF  
TITLE OF INVENTION: CARDIOVASCULAR AND OTHER DISEASES  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:

Query Match 12.9%; Score 77; DB 1; Length 14;  
Best Local Similarity 85.7%; Pred. No. 0.002;  
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 98 AGCKNFYWKGTSC 111  
Db 1 AGCKNFFWKFTSC 14



MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 502 or 55SX  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/286,748B  
FILING DATE: August 5, 1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: J. Peter Fasse  
REGISTRATION NUMBER: 32,983  
REFERENCE/DOCKET NUMBER: 04547/013001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-286-748B-13

Query Match 12.9%; Score 77; DB 1; Length 14;  
Best Local Similarity 85.7%; Pred. No. 0.002;  
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 98 AGCRNFWKGTSC 111  
|||||:|||||  
Db 1 AGCRNFWKGTSC 14

Search completed: March 21, 2003, 11:42:29  
Job time : 21.9614 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2003, 11:38:15 ; Search time 42.0773 Seconds  
(without alignments)  
421.131 Million cell updates/sec

US-09-727-739B-17  
Title: . . .  
Perfect score: 450  
Sequence: 1 MRVSIQHCAALLGLALATC.....RWRPRKVKGPQKAKEDLER 86

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 s

Database : SPTREMBL\_21: \*

- ```

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

```

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query |     | Length | DB     | ID     | Description  |
|------------|-------|-------|-----|--------|--------|--------|--------------|
|            |       | Match |     |        |        |        |              |
| 1          | 136   | 30.2  | 120 | 13     | Q90Y39 | Q90Y39 | catostomus   |
| 2          | 112   | 24.9  | 115 | 13     | Q90Y43 | Q90Y43 | osteoGLOSSu  |
| 3          | 108   | 24.0  | 114 | 13     | Q90Y42 | Q90Y42 | pantodon bu  |
| 4          | 93.5  | 20.8  | 114 | 13     | Q90Y41 | Q90Y41 | gnathonemus  |
| 5          | 88.5  | 19.7  | 116 | 13     | Q90XE1 | Q90XE1 | acipenser t  |
| 6          | 80.5  | 17.9  | 114 | 13     | Q90Y40 | Q90Y40 | chitala chi  |
| 7          | 78.5  | 17.4  | 603 | 16     | Q8XXA0 | Q8XXA0 | raistonia s  |
| 8          | 76    | 16.9  | 251 | 16     | Q98HG9 | Q98HG9 | rhizobium l  |
| 9          | 68.5  | 15.2  | 435 | 16     | Q9PAM4 | Q9PAM4 | xyliella fas |
| 10         | 68    | 15.1  | 421 | 10     | Q93W36 | Q93W36 | arabidopsis  |
| 11         | 68    | 15.1  | 578 | 10     | Q8S8B0 | Q8S8B0 | arabidopsis  |
| 12         | 68    | 15.1  | 755 | 10     | Q8S896 | Q8S896 | arabidopsis  |
| 13         | 68    | 15.1  | 758 | 10     | Q82440 | Q82440 | arabidopsis  |
| 14         | 68    | 15.1  | 772 | 10     | Q92SP5 | Q92SP5 | arabidopsis  |
| 15         | 67.5  | 15.0  | 197 | 16     | Q8ZMR3 | Q8ZMR3 | salmonella s |
| 16         | 67.5  | 15.0  | 568 | 16     | Q8YLB0 | Q8YLB0 | raistonia s  |

## ALIGNMENTS

## RESULT 1

|        |                                                                                                                                                                                    |      |          |
|--------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|----------|
| Q90Y39 | PRELIMINARY;                                                                                                                                                                       | PRT; | 120 AA.. |
| ID     | Q90Y39                                                                                                                                                                             |      |          |
| AC     | Q90Y39;                                                                                                                                                                            |      |          |
| DT     | 01-DEC-2001 (TREMBLrel. 19, Created)                                                                                                                                               |      |          |
| DT     | 01-DEC-2001 (TREMBLrel. 19, Last sequence update)                                                                                                                                  |      |          |
| DT     | 01-MAR-2002 (TREMBLrel. 20, Last annotation update)                                                                                                                                |      |          |
| DE     | Preprosomatostatin.                                                                                                                                                                |      |          |
| OS     | Catostomus commersoni (White sucker).                                                                                                                                              |      |          |
| OC     | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostei; Osteichthyes; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cyprinodontiformes; Catostomidae; Catostomus. |      |          |
| OC     | Catostomidae; Catostomus.                                                                                                                                                          |      |          |
| OX     | NCBI_TaxID=7971;                                                                                                                                                                   |      |          |
| RN     | [1]                                                                                                                                                                                |      |          |
| RP     | SEQUENCE FROM N.A.                                                                                                                                                                 |      |          |
| RA     | Al-Mahrouki A.A., Irwin D.M., Youson J.H.;                                                                                                                                         |      |          |
| RT     | "Molecular cloning and characterization of white sucker preprosomatostatin.";                                                                                                      |      |          |
| RT     | Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.                                                                                                                            |      |          |
| RRL    | EMBL; AF292654; AAR97071.1; -                                                                                                                                                      |      |          |
| DR     | InterPro; IPR004250; Somatostatin.                                                                                                                                                 |      |          |
| DR     | Pfam; PF03002; Somatostatin; 1.                                                                                                                                                    |      |          |
| DR     | SEQUENCE 120 AA; 13783 MW; 00828D35263E8805 CRC64;                                                                                                                                 |      |          |

RESULT 2  
Q90Y43



```
ID Q90Y43 PRELIMINARY; PRT; 115 AA.
AC Q90Y43;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Preprosomatostatin.
OS Osteoglossum bicirrhosum (silver arawana).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Osteoglossidae; Osteoglossum.
OX NCBI_TaxID=109271;
RN [1]
RP SEQUENCE FROM N.A.
RA Al-Mahrouki A.A., Irwin D.M., Youson J.H.;
RT "Characterization of variant somatostatin cDNAs from several
osteoglossomorphs: molecular identification and comparative
analysis.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF292650; AAK97067.1;
DR InterPro; IPR004250; Somatostatin.
DR Pfam; PF03002; Somatostatin; 1.
SQ SEQUENCE 115 AA; 12791 MW; D65FBD7C6F1E4E4D CRC64;

Query Match 24.9%; Score 112; DB 13; Length 115;
Best Local Similarity 56.5%; Pred. No. 3.6e-05;
Matches 26; Conservative 4; Mismatches 14; Indels 2; Gaps 1;

QY 1 MRVSIHCALALLGLALAI CSQGAASQPDLDLASRRLLQRAAL 46
Db 1 MKICQIHCTLVLLGLVLCPSAASQP--DLRYRSFLQRAHAAM 44

RESULT 3
Q90Y42
ID Q90Y42 PRELIMINARY; PRT; 114 AA.
AC Q90Y42;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Preprosomatostatin.
OS Pantodon buchholzi (Butterflyfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Pantodontidae; Pantodon.
OX NCBI_TaxID=8276;
RN [1]
RP SEQUENCE FROM N.A.
RA Al-Mahrouki A.A., Irwin D.M., Youson J.H.;
RT "Characterization of variant somatostatin cDNAs from several
osteoglossomorphs: molecular identification and comparative
analysis.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF292651; AAK97068.1;
DR InterPro; IPR004250; Somatostatin.
DR Pfam; PF03002; Somatostatin; 1.
SQ SEQUENCE 114 AA; 12352 MW; 7E3D44CB6A27B12F CRC64;

Query Match 24.0%; Score 108; DB 13; Length 114;
Best Local Similarity 50.0%; Pred. No. 0.0001;
Matches 23; Conservative 9; Mismatches 12; Indels 2; Gaps 1;

QY 1 MRVSIHCALALLGLALAI CSQGAASQPDLDLASRRLLQRAAL 46
Db 1 MKLCQVHCILALLGLVLCGSSSATQ--LDSRYRSLVQRAAAM 44

RESULT 4
Q90Y41
ID Q90Y41 PRELIMINARY; PRT; 114 AA.
AC Q90Y41;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
```

```
DE Preprosomatostatin.
OS Gnathepus petasii.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Mormyridae; Gnathepus.
OX NCBI_TaxID=42645;
RN [1]
RP SEQUENCE FROM N.A.
RA Al-Mahrouki A.A., Irwin D.M., Youson J.H.;
RT "Characterization of variant somatostatin cDNAs from several
osteoglossomorphs: molecular identification and comparative
analysis.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF292652; AAK97069.1;
DR InterPro; IPR004250; Somatostatin.
DR Pfam; PF03002; Somatostatin; 1.
SQ SEQUENCE 114 AA; 12494 MW; 454DA57A309CABF2 CRC64;

Query Match 20.8%; Score 93.5; DB 13; Length 114;
Best Local Similarity 56.8%; Pred. No. 0.0045;
Matches 25; Conservative 4; Mismatches 12; Indels 3; Gaps 1;

QY 1 MRVSIHCALALLGLALAI CSQGAASQPDLDLASRRLLQRAAL 44
Db 1 MLSSRIQCALALLSLALAVSSVSAAPS---DLKRLQLQRLSLA 41

RESULT 5
Q90XEL
ID Q90XEL PRELIMINARY; PRT; 116 AA.
AC Q90XEL;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Somatostatin.
OS Acipenser transmontanus (White sturgeon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
OC Acipenser.
OX NCBI_TaxID=7904;
RN [1]
RP SEQUENCE FROM N.A.
RA Trabucchi M., Tostivint H., Lihmann I., Sollars C., Vallarino M.,
Dores R.M., Vaudry H.;
RT "Polygenic expression of somatostatin in the sturgeon Acipenser
transmontanus: molecular cloning and distribution of the mRNAs
encoding two somatostatin precursors.";
RL J. Comp. Neurol. 0:0-0(2001).
DR EMBL; AF395849; AAL13248.1;
DR InterPro; IPR004250; Somatostatin.
DR Pfam; PF03002; Somatostatin; 1.
SQ SEQUENCE 116 AA; 12616 MW; 72E0C3FF6C80650F CRC64;

Query Match 19.7%; Score 88.5; DB 13; Length 116;
Best Local Similarity 54.3%; Pred. No. 0.017;
Matches 25; Conservative 5; Mismatches 11; Indels 5; Gaps 2;

QY 1 MRVSIHCALALLGLALAI CSQGAASQPDLDLASRRLLQRAAL 45
Db 1 MLSSRLQCALAFSLALAVSSVSAAPSPRL---RQLLQRTLAAS 42

RESULT 6
Q90Y40
ID Q90Y40 PRELIMINARY; PRT; 114 AA.
AC Q90Y40;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Preprosomatostatin.
OS Chitala chitala (clown knife-fish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```



OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;  
OC Osteoglossiformes; Notopteridae; Chitola.  
OX NCBI\_TaxID=112163;  
RN [1]  
RP SEQUENCE FROM N.A.

RA Al-Mahrouki A.A., Irwin D.M., Youson J.H.;  
RT "Characterization of variant somatostatin cDNAs from several  
RT osteoglossomorphs: molecular identification and comparative  
RT analysis";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF292653; AAK97070.1;  
DR InterPro; IPR004250; Somatostatin.  
DR Pfam; PF03002; Somatostatin; 1.  
SQ SEQUENCE 114 AA; 12561 MW; 4E3C32F58E34F971 CRC64;

Query Match 17.9%; Score 80.5; DB 13; Length 114;  
Best Local Similarity 50.0%; Pred. No. 0.13;  
Matches 22; Conservative 6; Mismatches 13; Indels 3; Gaps 1;

QY 1 MRVSIHCALALGLALAIICSGAASQPDLDLASRRLLQALAA 44  
Db 1 MLSTRIQCALALLSLALPVSSVYAAPS--DLKRLQLQRSIIA 41

RESULT 7

Q8XXAO PRELIMINARY; PRT; 603 AA.

AC Q8XXAO;  
DT 01-MAR-2002 (TReMBLrel. 20, Created)  
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)  
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
DE Probable DNA primase protein (EC 2.7.7.-).  
GN DNAG OR RSC2216 OR RS01383.  
OS Ralstonia solanacearum (Pseudomonas solanacearum).  
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;  
OC Ralstonia.  
OX NCBI\_TaxID=305;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GM11000;  
RX MEDLINE=21681879; PubMed=11823852;  
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,  
RA Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,  
RA Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Denange N.,  
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,  
RA Siguiet P., Thebault P., Whalen M., Wincker P., Levy M.,  
RA Weissenbach J., Boucher C.A.;  
RT "Genome sequence of the plant pathogen Ralstonia solanacearum";  
RL Nature 415:497-502(2002).  
DR EMBL; AL646068; CAD15923.1;  
DR InterPro; IPR002936; DNAPrim\_toprim.  
DR InterPro; IPR002694; Znf\_CHC2.  
DR Pfam; PF01751; Toprim; 1.  
DR Pfam; PF01807; zf-CHC2; 1.  
DR ProDom; PD002988; Znf\_CHC2; 1.  
DR SMART; SM00493; TOPRIM; 1.  
DR SMART; SM00400; Znf\_CHCC; 1.  
KW Transferrase; Nucleotidyltransferase; Complete proteome.  
SQ SEQUENCE 603 AA; 66305 MW; 5A3A29A1674E541C CRC64;

Query Match 17.4%; Score 78.5; DB 16; Length 603;  
Best Local Similarity 38.6%; Pred. No. 1.2;  
Matches 22; Conservative 3; Mismatches 15; Indels 17; Gaps 1;

QY 9 ALALLGLALAIICSGAASQP-----DLDLASRRLLQALAAALPH 48  
Db 278 ALAQLGFANAVATLGTACTPVHVQKLLRQVDVIFSFDDGDAAGRAARRALEACLPH 334

RESULT 8

Q98HG9 PRELIMINARY; PRT; 251 AA.

AC Q98HG9;

DT 01-OCT-2001 (TReMBLrel. 18, Created)  
DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)  
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
DE Hypothetical protein ml12874.  
GN MLL2874.  
OS Rhizobium loti (Mesorhizobium loti).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Phyllobacteriaceae; Mesorhizobium.  
OX NCBI\_TaxID=381;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=MAFF303099; PubMed=11214968;  
RX MEDLINE=21082930; PubMed=11214968;  
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,  
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,  
RA Takeuchi C., Yamada M., Tabata S.;  
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
RT Mesorhizobium loti";  
RL DNA Res. 7:331-338(2000).  
DR EMBL; AP003000; BAB49897.1;  
DR InterPro; IPR001173; Glycos\_transf\_2.  
DR Pfam; PF00535; Glycos\_transf\_2; 1.  
KW Hypothetical protein; Complete proteome.

Query Match 16.9%; Score 76; DB 16; Length 251;  
Best Local Similarity 33.3%; Pred. No. 0.97;  
Matches 29; Conservative 15; Mismatches 23; Indels 20; Gaps 5;

QY 3 VSQIHICALALGLALAIICSGAASQPDLDLASRRLLQALAAALPHRSVGSERWRTFYPN 62  
Db 180 VSALKTREGLLALKIA--RVHYARGDLDAAREL--AVAAAAPKRR--SEAWR----- 226  
QY 63 CPCLRWRR-----PRKVGKPOLKAKEDL 84  
Db 227 --CLRYRLKLAVRRLSAPKIEQLSVL 251

RESULT 9

Q9PAM4 PRELIMINARY; PRT; 435 AA.

AC Q9PAM4;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
DE Hypothetical protein Xf2472.  
GN Xf2472.  
OS Xylella fastidiosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
OC Xylella.  
OX NCBI\_TaxID=2371;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=9A5C;  
RX MEDLINE=20365717; PubMed=10910347;  
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,  
RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,  
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,  
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,  
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,  
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,  
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,  
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,  
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kenper E.L., Kitajima J.P.,  
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,  
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,  
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,  
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,  
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,  
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,







RA Fraser C.M., Venter J.C.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC007293; AAM15483.1;  
KW Protease.  
FT NON\_TER 755 755  
SQ SEQUENCE 755 AA; 80844 MW; 2E9BE365795936DE CRC64;  
  
Query Match 15.1%; Score 68; DB 10; Length 755;  
Best Local Similarity 28.1%; Pred. No. 24;  
Matches 18; Conservative 13; Mismatches 25; Indels 8; Gaps 2;  
  
QY 14 GLALAICSGAASQPDLDLASRLLQALAA---ALPHRSGVSEWRWTFYPNCPCLRWRP 70  
Db 529 GVSVIAAYTGAVSPTNEQFDPRLFLNAISGTSMSCPHISGIAGLLKTRYP-----WSP 583  
QY 71 RKVK 74  
Db 584 AAIR 587  
  
RESULT 13  
O82440 PRELIMINARY; PRT; 758 AA.  
AC O82440;  
DT 01-NOV-1998 (TReMBLrel. 08, Created)  
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)  
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
DE Subtilisin-like protease (Fragment).  
GN AIR3.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=ROOT CULTURE;  
RX MEDLINE=99178779; PubMed=10080694;  
RA Neuteboom L.W., Ng J.M.Y., Kuiper M., Clifdesdale O.R.,  
RA Hooykaas P.J.J., van der Zaai B.J.;  
RT "Isolation and characterization of cDNA clones corresponding with  
mRNAs that accumulate during auxin-induced lateral root formation";  
RL Plant Mol. Biol. 39:273-287(1999).  
DR EMBL; AF055848; AAC62611.1;  
DR HSSP; Q45670; IDBI.  
DR InterPro; IPR003137; PA.  
DR InterPro; IPR000209; Peptidase\_S8.  
DR Pfam; PF02225; PA; 1.  
DR Pfam; PF00082; Peptidase\_S8; 2.  
DR PRINTS; PR00723; SUBTILISIN.  
DR PROSITE; PS00138; SUBTILASE\_SER; UNKNOWN\_1.  
KW Protease.  
FT NON\_TER 1  
SQ SEQUENCE 758 AA; 81233 MW; E1683D2AD6C0516C CRC64;  
  
Query Match 15.1%; Score 68; DB 10; Length 758;  
Best Local Similarity 28.1%; Pred. No. 24;  
Matches 18; Conservative 13; Mismatches 25; Indels 8; Gaps 2;  
  
QY 14 GLALAICSGAASQPDLDLASRLLQALAA---ALPHRSGVSEWRWTFYPNCPCLRWRP 70  
Db 515 GVSVIAAYTGAVSPTNEQFDPRLFLNAISGTSMSCPHISGIAGLLKTRYP-----WSP 569  
QY 71 RKVK 74  
Db 570 AAIR 573  
  
RESULT 14  
Q9ZSP5 PRELIMINARY; PRT; 772 AA.  
AC Q9ZSP5;  
DT 01-MAY-1999 (TReMBLrel. 10, Created)

DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
DE Subtilisin-like protease.  
GN AIR3.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RX MEDLINE=99246056; PubMed=10231025;  
RA Neuteboom L.W., Veth-Tello L.M., Clifdesdale O.R., Hooykaas P.J.,  
RA van der Zaai B.J.;  
RT "A novel subtilisin-like protease gene from Arabidopsis thaliana is  
expressed at sites of lateral root emergence";  
RL DNA Res. 6:13-19(1999).  
DR EMBL; AF098632; AAD12260.1;  
DR HSSP; Q45670; IDBI.  
DR InterPro; IPR003137; PA.  
DR InterPro; IPR000209; Peptidase\_S8.  
DR Pfam; PF02225; PA; 1.  
DR Pfam; PF00082; Peptidase\_S8; 2.  
DR PRINTS; PR00723; SUBTILISIN.  
DR PROSITE; PS00138; SUBTILASE\_SER; UNKNOWN\_1.  
KW Protease.  
SQ SEQUENCE 772 AA; 82873 MW; 75DCCA0DED63F47E CRC64;  
  
Query Match 15.1%; Score 68; DB 10; Length 772;  
Best Local Similarity 28.1%; Pred. No. 24;  
Matches 18; Conservative 13; Mismatches 25; Indels 8; Gaps 2;  
  
QY 14 GLALAICSGAASQPDLDLASRLLQALAA---ALPHRSGVSEWRWTFYPNCPCLRWRP 70  
Db 529 GVSVIAAYTGAVSPTNEQFDPRLFLNAISGTSMSCPHISGIAGLLKTRYP-----WSP 583  
QY 71 RKVK 74  
Db 584 AAIR 587  
  
RESULT 15  
Q8ZMR3 PRELIMINARY; PRT; 197 AA.  
AC Q8ZMR3;  
DT 01-MAR-2002 (TReMBLrel. 20, Created)  
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)  
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
DE Putative periplasmic protein.  
GN STM2741.  
OS Salmonella typhimurium.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Salmonella.  
OX NCBI\_TaxID=602;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LT2 / SCSC1412 / ATCC 700720;  
RX MEDLINE=21534948; PubMed=11677609;  
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,  
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,  
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
RA Ryan B., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,  
RA Waterston R., Wilson R.K.;  
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium  
LT2";  
RL Nature 413:852-856(2001).  
DR EMBL; AE008824; AAL21627.1;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 197 AA; 22326 MW; 7CA059444F962F66 CRC64;  
  
Query Match 15.0%; Score 67.5; DB 16; Length 197;  
Best Local Similarity 33.3%; Pred. No. 7;



|    | Matches | 30;                                                         | Conservative | 9;    | Mismatches | 38; | Indels | 13; | Gaps | 4; |
|----|---------|-------------------------------------------------------------|--------------|-------|------------|-----|--------|-----|------|----|
| QY | 7       | HCAIALLGLAIGCSGAASQPDLDSLRRLLQRALAAAALPHRSGVS--ERWRTEFYPN-C | 63           |       |            |     |        |     |      |    |
|    |         |                                                             |              | : : : | : : :      | :   |        | :   | :    | :  |
| Db | 8       | HLALLLGLD---ASQPANALPEEHQENRYILHDAISRTIKATGMVSAPANKRMFYADEM | 64           |       |            |     |        |     |      |    |
| QY | 64      | PCLRWR-----PRKVKGPQLKAKEDLER                                | 86           |       |            |     |        |     |      |    |
|    |         |                                                             | :            |       |            |     |        |     |      |    |
| Db | 65      | FALAWRLIDDELTPESEVKDKDRSLKAVMKLSR                           | 94           |       |            |     |        |     |      |    |

Search completed: March 21, 2003, 11:40:43  
Job time : 46.0773 secs







```
RESULT 2
SMS2_CARAU
ID SMS2_CARAU STANDARD; PRT; 120 AA.
AC Q9YGH4; Q9PTU2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin II precursor [Contains: {Tyr21,Gly24}somatostatin-28;
DE [Tyr7,Gly10]somatostatin-14].
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Carassius.
OX NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Lin X.-W., Peter R.E.;
RT "Cloning and characterization of cDNAs encoding preprosomatostatin-I
RT and -II from goldfish brain.";
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U60262; AAD09626.1;
DR InterPro; IPR004250; Somatostatin.
DR Pfam; PF03002; Somatostatin; 1.
KW Cleavage on pair of basic residues; Hormone; Signal; Multigene family.
FT SIGNAL 1 23 POTENTIAL.
FT PROPEP 24 92 POTENTIAL.
FT PEPTIDE 93 120 [TYR21, GLY24]SOMATOSTATIN-28.
FT PEPTIDE 107 120 [TYR7, GLY10]SOMATOSTATIN-14.
FT DISULFID 109 120 BY SIMILARITY.
FT CONFLICT 51 51 Q -> RW (IN REF. 2).
SEQUENCE 120 AA; 13723 MW; 98957D68011A651A CRC64;
Query Match. 26.4%; Score 119; DB 1; Length 120;
Best Local Similarity 57.8%; Pred. No. 1.3e-06;
Matches 26; Conservative 6; Mismatches 11; Indels 2; Gaps 1;
QY 1 MRVSIQHICALALLGLALAIICSGQAASQ--PDLDASRRLLQRALA 43
DB 1 MRLGELHCYTLALLGLSLVLCGRCAANSQLEPDLPFRHRLQALASA 45
RESULT 3
SMSA_CARAU
ID SMSA_CARAU STANDARD; PRT; 114 AA.
AC Q9YGH5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin IA precursor [Contains: Somatostatin-26; Somatostatin-
DE 14].
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Carassius.
OX NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Lin X.-W., Peter R.E.;
RT "Cloning and characterization of cDNAs encoding preprosomatostatin-I
RT and -II from goldfish brain.";
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U60262; AAD09626.1;
DR InterPro; IPR004250; Somatostatin.
DR Pfam; PF03002; Somatostatin; 1.
KW Cleavage on pair of basic residues; Hormone; Signal; Multigene family.
FT SIGNAL 1 23 POTENTIAL.
FT PROPEP 24 92 POTENTIAL.
FT PEPTIDE 93 120 [TYR21, GLY24]SOMATOSTATIN-28.
FT PEPTIDE 107 120 [TYR7, GLY10]SOMATOSTATIN-14.
FT DISULFID 109 120 BY SIMILARITY.
FT CONFLICT 51 51 Q -> RW (IN REF. 2).
SEQUENCE 120 AA; 13723 MW; 98957D68011A651A CRC64;
Query Match. 26.4%; Score 119; DB 1; Length 120;
Best Local Similarity 57.8%; Pred. No. 1.3e-06;
Matches 26; Conservative 6; Mismatches 11; Indels 2; Gaps 1;
QY 1 MRVSIQHICALALLGLALAIICSGQAASQ--PDLDASRRLLQRALA 43
DB 1 MRLGELHCYTLALLGLSLVLCGRCAANSQLEPDLPFRHRLQALASA 45
RESULT 4
SMS_CHICK
ID SMS_CHICK STANDARD; PRT; 116 AA.
AC P33094;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin precursor [Contains: Somatostatin-28; Somatostatin-14].
DE SST.
GN Gallus gallus (Chicken).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RA Nata K., Kobayashi T., Karahashi K., Kato S., Yamamoto H.,
RA Yonekura H., Okamoto H.;
RL Submitted (JUN-1991) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U40754; AAD09359.1;
DR InterPro; IPR004250; Somatostatin.
DR Pfam; PF03002; Somatostatin; 1.
KW Cleavage on pair of basic residues; Hormone; Signal; Multigene family.
FT SIGNAL 1 24 POTENTIAL.
FT PROPEP 25 88 POTENTIAL.
FT PEPTIDE 89 114 SOMATOSTATIN-26 (POTENTIAL).
FT PEPTIDE 101 114 SOMATOSTATIN-14.
FT DISULFID 103 114 BY SIMILARITY.
SEQUENCE 114 AA; 12574 MW; B5920015E2D272A4 CRC64;
Query Match. 20.3%; Score 91.5; DB 1; Length 114;
Best Local Similarity 54.8%; Pred. No. 0.0016;
Matches 23; Conservative 5; Mismatches 11; Indels 3; Gaps 1;
QY 1 MRVSIQHICALALLGLALAIICSGQAASQPDLDASRRLLQRAL 42
DB 1 MLESTRIQCALLALLSLALAVCSVSA--PTDAKLRLLQRLS 39
```







```
CC tumors and vasoactive intestinal peptide tumors.
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J00306; AAA60566.1; -
CC EMBL; M19318; AAA36908.1; -
CC PIR; A43614; RIHUS1.
CC PIR; A28968; A28968.
CC Genew; HGNC:11329; SST.
CC MIM; 182450; -
CC InterPro; IPR004250; Somatostatin.
CC Pfam; PF03002; Somatostatin; 1.
CC Cleavage on pair of basic residues; Hormone; Signal; Pharmaceutical.
FT SIGNAL 1 24
FT PROPEP 25 88
FT PEPTIDE 89 116 SOMATOSTATIN-28.
FT PEPTIDE 103 116 SOMATOSTATIN-14.
FT DISULFID 105 116
FT SEQUENCE 116 AA; 12735 MW; AB49BB89DC9DD8DA CRC64;
SQ
Query Match 17.7%; Score 79.5; DB 1; Length 116;
Best Local Similarity 34.9%; Pred. No. 0.037;
Matches 29; Conservative 13; Mismatches 28; Indels 13; Gaps 5;
QY 5 QIHCALALLGLALAI-CSQGAASQPDLDLASRRLLQRLAALPHRSGVSEWRRTFYPNC 63
Db 5 RLQCALAALSIVLALGCVTGAPSDPRL----RQFLQKSLAAA----AGKQELAKYFLAE- 55
QY 64 PCLWRPRKVKGPQLKAKEDLER 86
Db 56 --LLSEPNQTENDALE-PEDLSQ 75
RESULT 7
SMS1_RANR1 STANDARD; PRT; 115 AA.
AC P87384; Q9PSI8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin 1 precursor (PSS1) [Contains: Somatostatin-14 (S-1)
DE (SSS1)].
OS Rana ridibunda (Laughing frog) (Marsh frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranolide; Ranidae; Rana.
NCBI_TaxID=8406;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97057290; PubMed=8901629;
RA Tostivint H., Lihrmann I., Bucharles C., Vieau D., Coulouarn Y.,
RA Fournier A., Conlon J.M., Vaudry H.;
RT "Occurrence of two somatostatin variants in the frog brain:
RT characterization of the cDNAs, distribution of the mRNAs, and
RT receptor-binding affinities of the peptides."
RL proc. Natl. Acad. Sci. U.S.A. 93:12605-12610(1996).
RN [2]
RP SEQUENCE OF 102-115.
RC TISSUE=Brain;
RX MEDLINE=93038702; PubMed=1358069;
RA Vaudry H., Chartrel N., Conlon J.M.;
RT "Isolation of [Pro2,Met13]somatostatin-14 and somatostatin-14 from the
RT frog brain reveals the existence of a somatostatin gene family in a
RT tetrapod."
RL Blochem. Biophys. Res. Commun. 188:477-482(1992).
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
```

```
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U68136; AAC60093.1; -
CC InterPro; IPR004250; Somatostatin.
CC Pfam; PF03002; Somatostatin; 1.
CC Cleavage on pair of basic residues; Hormone; Multigene family; Signal.
FT SIGNAL 1 24 BY SIMILARITY.
FT PROPEP 25 99 BY SIMILARITY.
FT PEPTIDE 102 115 SOMATOSTATIN-14.
FT DISULFID 104 115 BY SIMILARITY.
FT SEQUENCE 115 AA; 12691 MW; 349756FEBA4BE213 CRC64;
SQ
Query Match 17.0%; Score 76.5; DB 1; Length 115;
Best Local Similarity 47.8%; Pred. No. 0.079;
Matches 22; Conservative 8; Mismatches 11; Indels 5; Gaps 2;
QY 1 MRVSQIHCALALLGLALAICSQGAA-SQPDLDLASRRLLQRLAALAA 45
Db 1 MQSCRVQCALTLLSLALAINISAAATDPRL---RQFLQKSLASA 42
RESULT 8
SMS1_PROAN STANDARD; PRT; 115 AA.
AC Q9W7F0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin 1 precursor (PSS1) [Contains: Somatostatin-27;
DE Somatostatin-14].
OS Protopterus annectens (African lungfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Dipnoi; Lepidosireniiformes; Protopteriidae; Protopterus.
NCBI_TaxID=7888;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99326690; PubMed=10398054;
RA Trabucchi M., Tostivint H., Lihrmann I., Jegou S., Vallarino M.,
RA Vaudry H.;
RT "Molecular cloning of the cDNAs and distribution of the mRNAs encoding
RT two somatostatin precursors in the African lungfish Protopterus
RT annectens."
RL J. Comp. Neurol. 410:643-652(1999).
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF126243; AAD39138.1; -
CC InterPro; IPR004250; Somatostatin.
CC Pfam; PF03002; Somatostatin; 1.
CC Cleavage on pair of basic residues; Hormone; Multigene family; Signal.
FT SIGNAL 1 24 POTENTIAL.
FT PROPEP 25 88 POTENTIAL.
FT PEPTIDE 89 115 SOMATOSTATIN-27 (POTENTIAL).
FT PEPTIDE 102 115 SOMATOSTATIN-14.
```



FT DISULFID 104 115 BY SIMILARITY.  
SQ SEQUENCE 115 AA; 12600 MW; B0CEFI603FEAF09 CRC64;

Query Match 16.8%; Score 75.5; DB 1; Length 115;  
Best Local Similarity 48.8%; Pred. No. 0.1;  
Matches 20; Conservative 4; Mismatches 14; Indels 3; Gaps 1;

QY 5 QIHCALALLGLALAICSGQASQPDLDLASRRLLQRALAAA 45  
Db 5 RFQCALVLLSLAVFESKVSAAPS---DLRLRQLLQRLSAAA 42

RESULT 9  
SMS\_SHEEP

ID SMS\_SHEEP STANDARD; PRT; 116 AA.  
AC O46588;  
OT 15-DEC-1998 (Rel. 37, Created)  
JT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Somatostatin precursor [Contains: Somatostatin-28; Somatostatin-14].  
GN SST.  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Ile de France;  
RX MEDLINE-99094691; PubMed-9880082;  
RA Bruneau G., Tillet Y.;  
RT "Localization of the preprosomatostatin-mRNA by in situ hybridization  
in the ewe hypothalamus."  
RL Peptides 19:1749-1758(1998).  
CC -!- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
CC EMBL; AF031488; AAC04697.1;  
DR EMBL; Y15267; CAA75556.1;  
DR InterPro; IPR004250; Somatostatin.  
DR Pfam; PF03002; Somatostatin; 1.  
KW Cleavage on pair of basic residues; Hormone; Signal.  
FT SIGNAL 1 24 BY SIMILARITY.  
FT PROPEP 25 88 BY SIMILARITY.  
FT PEPTIDE 89 116 SOMATOSTATIN-28.  
FT PEPTIDE 103 116 SOMATOSTATIN-14.  
FT DISULFID 105 116 BY SIMILARITY.  
SQ SEQUENCE 116 AA; 12689 MW; C18FI7E31A3718DE CRC64;

Query Match 15.4%; Score 69.5; DB 1; Length 116;  
Best Local Similarity 33.7%; Pred. No. 0.49;  
Matches 28; Conservative 13; Mismatches 29; Indels 13; Gaps 5;

QY 5 QIHCALALLGLALAICSGQASQPDLDLASRRLLQRALAAAALPHRSGVSEWRTFFYPNC 63  
Db 5 RLQCALAALSIVLAIGVGTGAPSDPRL----RQFLQKSLAAA---AGKQLAKYFLAE- 55

QY 64 PCRLWRPRKVKGPQLKAKEDLER 86

Db 56 --LLSEPNQNDALAE-PEDLSQ 75

RESULT 10  
SMS2\_LOPAM

ID SMS2\_LOPAM STANDARD; PRT; 125 AA.  
AC P01170; O91066;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Somatostatin II precursor [Contains: [Tyr7,Gly10]somatostatin-14].  
OS Lophius americanus (American goosefish) (Anglerfish)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Paracanthopterygii; Lophiiformes; Lophiidae; Lophius.  
OX NCBI\_TaxID=8073;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-81052423; PubMed-6107860;  
RA Hobart P.M., Crawford R., Shen L., Pictet R., Rutter W.J.;  
RT "Cloning and sequence analysis of cDNAs encoding two distinct  
somatostatin precursors found in the endocrine pancreas of  
anglerfish."  
RL Nature 288:137-141(1980).  
RN [2]  
RP PARTIAL SEQUENCE, AND HYDROXYLATION.  
RX MEDLINE-87308304; PubMed-2887572;  
RA Andrews P.C., Nichols R., Dixon J.E.;  
RT "Post-translational processing of preprosomatostatin-II examined  
using fast atom bombardment mass spectrometry."  
RL J. Biol. Chem. 262:12692-12699(1987).  
CC -!- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- MISCELLANEOUS: SOMATOSTATIN II MAY HAVE A DIFFERENT DEGREE OF  
ACTIVITY OR A DIFFERENT TYPE OF TARGET CELL FROM SOMATOSTATIN I.  
CC -!- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
CC EMBL; V00641; CAA23987.1;  
DR PIR; A01434; RIAFS2.  
DR PIR; A27376; A27376.  
DR InterPro; IPR004250; Somatostatin.  
DR Pfam; PF03002; Somatostatin; 1.  
KW Cleavage on pair of basic residues; Hormone; Signal; Hydroxylation;  
KW Multigene family.  
FT SIGNAL 1 24 POTENTIAL.  
FT PROPEP 25 109  
FT PEPTIDE 112 125 [TYR7, GLY10]SOMATOSTATIN-14.  
FT DISULFID 114 125  
FT MOD.RES 120 120  
FT CONFLICT 77 78  
FT CONFLICT 90 90  
FT CONFLICT 90 90  
SQ SEQUENCE 125 AA; 14052 MW; 5E14605D7B9A46FE CRC64;

Query Match 15.4%; Score 69.5; DB 1; Length 125;  
Best Local Similarity 36.1%; Pred. No. 0.53;  
Matches 22; Conservative 7; Mismatches 17; Indels 15; Gaps 3;

QY 6 IHCALALLGLALAICSGQASQ-----PDLASRR--LLQRALAAAALPHRSGVSR 55  
Db 4 IRCPAILALLALVLCGFSVSSQLDREQSDNQDLDLELRQHWLLERARSAGL-----LSQE 58

QY 56 W 56

Db 59 W 59

RESULT 11

SMS\_BOVIN  
ID SMS\_BOVIN STANDARD; PRT; 116 AA.  
AC P26917;







nitric oxide synthase.";  
J. Biol. Chem. 267:15274-15276(1992).  
[4]  
MYRISTOYLATION.  
MEDLINE-93231982; PubMed-7682550;  
Busconi L., Michel T.;  
"Endothelial nitric oxide synthase: N-terminal myristoylation  
determines subcellular localization.";  
J. Biol. Chem. 268:8410-8413(1993).  
[5]  
PALMITOYLATION.  
MEDLINE-96102197; PubMed-8524847;  
Robinson L.J., Michel T.;  
"Mutagenesis of palmitoylation sites in endothelial nitric oxide  
synthase identifies a novel motif for dual acylation and subcellular  
targeting.";  
Proc. Natl. Acad. Sci. U.S.A. 92:11776-11780(1995).  
[6]  
X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 66-481.  
MEDLINE-99091052; PubMed-9875848;  
Raman C.S., Li H., Martasek P., Kral V., Masters B.S.S., Poulos T.L.;  
"Crystal structure of constitutive endothelial nitric oxide synthase:  
a paradigm for pterin function involving a novel metal center.";  
Cell 95:939-950(1998).  
[7]  
X-RAY CRYSTALLOGRAPHY (1.86 ANGSTROMS) OF 66-481.  
MEDLINE-20503854; PubMed-11051558;  
Li H., Raman C.S., Martasek P., Kral V., Masters B.S.S., Poulos T.L.;  
"Mapping the active site polarity in structures of endothelial  
nitric oxide synthase heme domain complexed with isothioureas.";  
J. Inorg. Biochem. 81:133-139(2000).  
[8]  
X-RAY CRYSTALLOGRAPHY (1.93 ANGSTROMS).  
MEDLINE-21229525; PubMed-11331003;  
Li H., Raman C.S., Martasek P., Masters B.S.S., Poulos T.L.;  
"Crystallographic studies on endothelial nitric oxide synthase  
complexed with nitric oxide and mechanism-based inhibitors.";  
Biochemistry 40:5399-5406(2001).  
[9]  
X-RAY CRYSTALLOGRAPHY (1.65 ANGSTROMS).  
MEDLINE-21552923; PubMed-11695891;  
Raman C.S., Li H., Martasek P., Southern G., Masters B.S.S.,  
Poulos T.L.;  
"Crystal structure of nitric oxide synthase bound to nitro indazole  
reveals a novel inactivation mechanism.";  
Biochemistry 40:13448-13455(2001).  
[10]  
X-RAY CRYSTALLOGRAPHY (1.93 ANGSTROMS).  
MEDLINE-21336567; PubMed-11331290;  
Raman C.S., Li H., Martasek P., Babu B.R., Griffith O.W., Southern G.,  
Masters B.S.S., Poulos T.L.;  
"Implications for isoform-selective inhibitor design derived from the  
binding mode of bulky isothioureas to the heme domain of endothelial  
nitric-oxide synthase.";  
J. Biol. Chem. 276:26486-26491(2001).  
-1- FUNCTION: PRODUCES NITRIC OXIDE (NO) WHICH IS IMPLICATED IN  
VASCULAR SMOOTH MUSCLE RELAXATION THROUGH A CGMP-MEDIATED SIGNAL  
TRANSDUCTION PATHWAY. NO MEDIATES VASCULAR ENDOTHELIAL GROWTH  
FACTOR (VEGF)-INDUCED ANGIOGENESIS IN CORONARY VESSELS AND  
PROMOTES BLOOD CLOTTING THROUGH THE ACTIVATION OF PLATELETS.  
-1- CATALYTIC ACTIVITY: L-arginine + N NADPH + M O(2) -> citrulline +  
nitric oxide + N NADP(+).  
-1- COFACTOR: HEME. BINDS ONE MOLE EACH OF FAD AND FMN. ALSO REQUIRES  
TETRAHYDROBIOTERIN (BH4) WHICH MAY STABILIZE THE DIMERIC FORM OF  
THE ENZYME.  
-1- ENZYME REGULATION: STIMULATED BY CALCIUM/CALMODULIN.  
-1- SUBUNIT: HOMODIMER.  
-1- SIMILARITY: BELONGS TO THE NOS FAMILY.  
-----  
THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))  
-----  
CC EMBL; M99057; AAA30667.1; -  
CC EMBL; M89952; AAA30494.1; -  
CC EMBL; M95674; AAA30669.1; -  
CC PDB; 1NSE; 18-MAY-99.  
CC PDB; 2NSE; 25-MAY-99.  
CC PDB; 3NSE; 18-MAY-99.  
CC PDB; 4NSE; 18-MAY-99.  
CC PDB; 8NSE; 21-NOV-01.  
CC PDB; 9NSE; 25-OCT-00.  
CC PDB; 1DIW; 25-OCT-00.  
CC PDB; 1ED4; 25-OCT-00.  
CC PDB; 1DM6; 13-DEC-00.  
CC PDB; 1DM7; 13-DEC-00.  
CC PDB; 1DM8; 13-DEC-00.  
CC PDB; 1DMI; 20-DEC-00.  
CC PDB; 1DMJ; 20-DEC-00.  
CC PDB; 1ED5; 31-JAN-01.  
CC PDB; 1ED6; 31-JAN-01.  
CC PDB; 1FOI; 20-JUL-01.  
CC PDB; 1FOL; 20-JUL-01.  
CC PDB; 1FOO; 20-JUL-01.  
CC PDB; 1FOP; 20-JUL-01.  
CC PDB; 1DIV; 25-JUL-01.  
CC PDB; 1DIX; 25-JUL-01.  
CC PDB; 1DIY; 25-JUL-01.  
CC PDB; 1DOC; 21-NOV-01.  
CC PDB; 1DOO; 21-NOV-01.  
CC PDB; 1FOJ; 16-NOV-01.  
CC InterPro; IPR003097; FAD\_binding.  
CC InterPro; IPR001709; FPN\_cyt\_redctse.  
CC InterPro; IPR001094; Flavodoxin\_like.  
CC InterPro; IPR001226; Flavodoxin.  
CC InterPro; IPR004030; NO\_synthase.  
CC InterPro; IPR001433; Oxred\_FAD/NAD(P).  
CC Pfam; PF00175; NAD\_binding; 1.  
CC Pfam; PF00258; flavodoxin; 1.  
CC Pfam; PF00667; FAD\_binding; 1.  
CC Pfam; PF02898; NO\_synthase; 1.  
CC PRINTS; PR00369; FLAVODOXIN.  
CC PRINTS; PR00371; FPNCR.  
CC PROSITE; PS60001; NOS; 1.  
KW Oxidoreductase; NADP; FAD; FMN; Calmodulin-binding; Myristate;  
KW Lipoprotein; Palmitate; Phosphorylation; Calcium-binding; Heme;  
KW Zinc; Metal-binding; Multigene family; 3D-structure.  
FT INIT\_MET 0 0  
FT BINDING 185 185 HEME.  
FT DOMAIN 491 511 CALMODULIN-BINDING (POTENTIAL).  
FT NP\_BIND 650 681 FMN (PYRIMIDINE PART) (BY SIMILARITY).  
FT NP\_BIND 794 805 FAD (ADP PART) (BY SIMILARITY).  
FT NP\_BIND 936 946 FAD (FLAVIN PART) (BY SIMILARITY).  
FT NP\_BIND 1011 1029 NADP (RIBOSE PART) (BY SIMILARITY).  
FT NP\_BIND 1109 1124 NADP (ADP PART) (BY SIMILARITY).  
FT LIPID 1 1 MYRISTATE.  
FT LIPID 14 14 PALMITATE.  
FT LIPID 25 25 PALMITATE.  
FT METAL 95 95 ZINC.  
FT METAL 100 100 ZINC.  
FT MOD\_RES 142 142 PHOSPHORYLATION (BY PKA).  
FT CONFLICT 99 99 C -> R (IN REF. 3).  
FT CONFLICT 164 164 Y -> I (IN REF. 3).  
FT CONFLICT 317 327 EHPTLEWFAAL -> GAPHTGVVRGP (IN REF. 3).  
FT CONFLICT 454 454 S -> Y (IN REF. 3).  
FT CONFLICT 458 458 T -> P (IN REF. 3).  
FT CONFLICT 740 740 T -> A (IN REF. 3).  
FT CONFLICT 803 804 CP -> V (IN REF. 3).  
FT CONFLICT 856 856 L -> V (IN REF. 3).  
FT CONFLICT 906 907 WF -> LV (IN REF. 3).  
FT CONFLICT 1041 1041 A -> H (IN REF. 3).  
SQ SEQUENCE 1204 AA; 133155 MW; D017210062ABE4B0 CRC64;



Query Match 14.9%; Score 67; DB 1; Length 1204;  
Best Local Similarity 35.5%; Pred. No. 10;  
Matches 27; Conservative 4; Mismatches 25; Indels 20; Gaps 5;  
QY 8. CALALLGLALAIC-SQGAAS---QPDLDLASRRLLQALAAALPHRSVGSERWRTFYPNC 63  
Db 14 CGLG-LGLGLGCGKQGPASPAEP-----SRAPAPATPHAPDHSPA-----PNS 57  
QY 64 PCLRWRRPRKVKGPQLK 79  
Db 58 PTLTRPPEGPKFPRVK 73

RESULT 14  
PT22\_SACBA STANDARD; PRT; 260 AA.  
AC O13374;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE PET122 protein, mitochondrial precursor.  
GN PET122.  
Saccharomyces bayanus (Yeast) (Saccharomyces uvarum).  
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
NCBI\_TaxID=4931;  
[1]  
SEQUENCE FROM N.A.  
RC STRAIN=NRRL-Y-12624;  
RA Costanzo M.C., Fox T.D.;  
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: REQUIRED FOR EXPRESSION OF THE MITOCHONDRIAL GENE FOR  
CYTOCHROME C OXIDASE SUBUNIT III (COX3). PET122 SEEMS TO WORKS BY  
DIRECTLY INTERACTING WITH THE SMALL RIBOSOMAL SUBUNIT TO PROMOTE  
TRANSLATION INITIATION ON THE COXIII MRNA (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL; INNER MEMBRANE ASSOCIATED (BY  
SIMILARITY).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF026394; AAB82600.1; --  
KW Translation regulation; Activator; Transit peptide; Mitochondrion;  
FT TRANSIT 1 8 MITOCHONDRION (BY SIMILARITY).  
CHAIN 9 260 PET122 PROTEIN.  
SEQUENCE 260 AA; 29016 MW; 80AFEC2A7238262F CRC64;

Query Match 14.7%; Score 66; DB 1; Length 260;  
Best Local Similarity 27.0%; Pred. No. 2:8;  
Matches 17; Conservative 10; Mismatches 30; Indels 6; Gaps 1;  
QY 13 LGLALICSGAASQPDLDLASRRLLQALAAALPHRSVGSERWRTF-----YPNCPCL 66  
Db 108 MGLGRGTSASASAEAAQAEFELRRVKVEAFARGTMHSTALSEKWKVFLQEMDTLPGQPPL 167  
QY 67 RWR 69  
Db 168 RLR 170

RESULT 15  
SMS\_MOUSE STANDARD; PRT; 116 AA.  
ID SMS\_MOUSE  
AC P01167;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Somatostatin precursor [Contains: Antrin; Somatostatin-28;  
DE Somatostatin-14].  
GN SST OR SMST.  
OS Mus musculus (Mouse), and  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090, 10116;  
RN [1].  
RP SEQUENCE FROM N.A.  
RC SPECIES=Rat;  
RX MEDLINE=85006903; PubMed=6148343;  
RA Tavianini M.A., Hayes T.E., Magazin M.D., Minth C.D., Dixon J.E.;  
RT "Isolation, characterization, and DNA sequence of the rat  
somatostatin gene."  
RL J. Biol. Chem. 259:11798-11803(1984).  
RN [2].  
RP SEQUENCE FROM N.A.  
RC SPECIES=Rat;  
RX MEDLINE=83238516; PubMed=6134734;  
RA Argos P., Taylor W.L., Minth C.D., Dixon J.E.;  
RT "Nucleotide and amino acid sequence comparisons of  
preprosomatostatins."  
RL J. Biol. Chem. 258:8788-8793(1983).  
RN [3].  
RP SEQUENCE FROM N.A.  
RC SPECIES=Rat;  
RX MEDLINE=83213516; PubMed=6133871;  
RA Goodman R.H., Aron D.C., Roos B.A.;  
RT "Rat pre-prosomatostatin. Structure and processing by microsomal  
membranes."  
RL J. Biol. Chem. 258:5570-5573(1983).  
RN [4].  
RP SEQUENCE FROM N.A.  
RC SPECIES=Rat;  
RX MEDLINE=85303584; PubMed=2863939;  
RA Goodman R.H., Montminy M.R., Low M.J., Habener J.F.;  
RT "Biosynthesis of rat preprosomatostatin."  
RL Adv. Exp. Med. Biol. 188:31-47(1985).  
RN [5].  
RP SEQUENCE FROM N.A.  
RC SPECIES=Rat;  
RX MEDLINE=84221954; PubMed=6145156;  
RA Montminy M.R., Goodman R.H., Horovitch S.J., Habener J.F.;  
RT "Primary structure of the gene encoding rat preprosomatostatin."  
RL Proc. Natl. Acad. Sci. U.S.A. 81:3337-3340(1984).  
RN [6].  
RP SEQUENCE OF 38-116 FROM N.A.  
RC SPECIES=Rat;  
RX MEDLINE=82120034; PubMed=6120163;  
RA Goodman R.H., Jacobs J.W., Dee P.C., Habener J.F.;  
RT "Somatostatin-28 encoded in a cloned cDNA obtained from a rat  
medullary thyroid carcinoma."  
RL J. Biol. Chem. 257:1156-1159(1982).  
RN [7].  
RP SEQUENCE OF 25-34.  
RC SPECIES=Rat; STRAIN=Sprague-Dawley;  
RX MEDLINE=88070564; PubMed=2891188;  
RA Benoit R., Ling N., Esch F.;  
RT "A new prosomatostatin-derived peptide reveals a pattern for  
prohormone cleavage at monobasic sites."  
RL Science 238:1126-1129(1987).  
RN [8].  
RP SEQUENCE FROM N.A.  
RC SPECIES=Mouse; TISSUE=Brain;  
RX MEDLINE=90206793; PubMed=1969620;  
RA Fuhrmann G., Heilig R., Kempf J., Ebel A.;  
RT "Nucleotide sequence of the mouse preprosomatostatin gene."  
RL Nucleic Acids Res. 18:1287-1287(1990).  
CC -!- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.



-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL; K02248; AAA42161.1; -  
DR EMBL; V01271; CAA24579.1; -  
DR EMBL; J00787; AAA42164.1; -  
DR EMBL; M25890; AAA42167.1; -  
DR EMBL; J00788; AAA42162.1; -  
DR EMBL; X51468; CAA35831.1; -  
DR PIR; A20983; RIRTS1  
DR PIR; S08416; S08416.  
DR MGD; MGI:98326; Smst.  
DR InterPro; IPR004250; Somatostatin.  
DR Pfam; PF03002; Somatostatin; 1.  
KW Cleavage on pair of basic residues; Hormone; Signal.  
FT SIGNAL 1 24  
FT PEPTIDE 25 34 ANTRIN.  
FT PROPEP 35 88  
FT PEPTIDE 89 116 SOMATOSTATIN-28.  
FT PEPTIDE 103 116 SOMATOSTATIN-14.  
FT DISULFID 105 116  
FT CONFLICT 43 43 T -> Y (IN REF. 5).  
FT CONFLICT 79 79 Q -> H (IN REF. 6).  
SQ SEQUENCE 116 AA; 12745 MW; D48B5454C4490375 CRC64;

Query Match 14.6%; Score 65.5; DB 1; Length 116;  
Best Local Similarity 45.2%; Pred. No. 1.4;  
Matches 19; Conservative 7; Mismatches 11; Indels 5; Caps 2;  
QY 5 QIHCALALLGLAICS-QGAASQPDLDLASRRLLORALAAA 45  
Db 5 RLQCALAALCIVLALGGVTGAPSDPRL-----RQFLQKSLAAA 42

Search completed: March 21, 2003, 11:38:40  
Job time : 13.3348 secs



GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2003, 11:38:56 ; Search time 18.4549 Seconds  
(without alignments)  
447.986 Million cell updates/sec

Title: US-09-727-739B-17  
Perfect score: 450  
Sequence: 1 MRVSIQHCALALLGLALAIC.....RWRPRKVGPKLAKEDLER 86  
Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5.

Searched: 283224 seqs, 96134422 residues 283224  
Total number of hits satisfying chosen parameters:  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_73:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID    | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 347.5 | 77.2        | 115    | 2 I51064 | somatostatin II pr |
| 2          | 83.5  | 18.6        | 116    | 1 S20630 | somatostatin precu |
| 3          | 80.5  | 17.9        | 114    | 1 RIIDS1 | somatostatin-14 pr |
| 4          | 80.5  | 17.9        | 114    | 2 I50798 | preprosomatostatin |
| 5          | 79.5  | 17.7        | 116    | 1 RIHUS1 | somatostatin I pre |
| 6          | 79.5  | 17.7        | 116    | 1 A28968 | somatostatin I pre |
| 7          | 76.5  | 17.0        | 115    | 2 JC6166 | somatostatin-14 pr |
| 8          | 69.5  | 15.4        | 125    | 1 RIAFS2 | somatostatin II pr |
| 9          | 68.5  | 15.2        | 116    | 1 RIBOS1 | somatostatin precu |
| 10         | 68.5  | 15.2        | 435    | 2 AB2554 | conserved hypothet |
| 11         | 68    | 15.1        | 758    | 2 T51335 | subtilisin-like pr |
| 12         | 67    | 14.9        | 120    | 2 F83244 | hypothetical prote |
| 13         | 67    | 14.9        | 1205   | 1 A38943 | nitric-oxide synth |
| 14         | 66.5  | 14.8        | 259    | 2 H81009 | exodeoxyribonuclea |
| 15         | 66.5  | 14.8        | 571    | 2 T43456 | hypothetical prote |
| 16         | 65.5  | 14.6        | 116    | 1 RIRTS1 | somatostatin precu |
| 17         | 65.5  | 14.6        | 116    | 1 RIMSS1 | somatostatin precu |
| 18         | 65.5  | 14.6        | 201    | 2 S55000 | alpha-chitin bindi |
| 19         | 64.5  | 14.3        | 1305   | 2 T31096 | cyclin G-associate |
| 20         | 63    | 14.0        | 431    | 2 F84375 | N-ethylammine ch   |
| 21         | 62.5  | 13.9        | 531    | 2 B83082 | probable binding p |
| 22         | 62    | 13.8        | 308    | 1 S76941 | carbamate kinase ( |
| 23         | 62    | 13.8        | 472    | 2 T36529 | probable two-compo |
| 24         | 62    | 13.8        | 501    | 2 AF0346 | probable GntR-famI |
| 25         | 61.5  | 13.7        | 256    | 2 AG0676 | probable ATP/GTP-b |
| 26         | 61.5  | 13.7        | 327    | 2 T34666 | probable regulator |
| 27         | 61.5  | 13.7        | 337    | 2 C83059 | conserved hypothet |
| 28         | 61.5  | 13.7        | 398    | 2 C81141 | hypothetical prote |
| 29         | 61.5  | 13.7        | 706    | 2 T01351 | subtilisin-like pr |

|    |      |      |      |          |                    |
|----|------|------|------|----------|--------------------|
| 30 | 61   | 13.6 | 238  | 2 H83559 | hypothetical prote |
| 31 | 61   | 13.6 | 375  | 2 G81878 | probable lipoprote |
| 32 | 61   | 13.6 | 519  | 2 D70695 | hypothetical prote |
| 33 | 60.5 | 13.4 | 273  | 2 E75476 | hypothetical prote |
| 34 | 60.5 | 13.4 | 321  | 2 G90894 | hypothetical prote |
| 35 | 60.5 | 13.4 | 321  | 2 B85723 | hypothetical prote |
| 36 | 60.5 | 13.4 | 469  | 2 S16359 | adenylate cyclase  |
| 37 | 60.5 | 13.4 | 474  | 2 C75625 | hypothetical prote |
| 38 | 60.5 | 13.4 | 1235 | 2 D32433 | VSG expression sit |
| 39 | 60   | 13.3 | 321  | 2 F75287 | probable DNA polym |
| 40 | 60   | 13.3 | 433  | 2 H86597 | flagellum-specific |
| 41 | 60   | 13.3 | 433  | 2 B72025 | flagellum-specific |
| 42 | 60   | 13.3 | 433  | 2 A81513 | virulence ATPase,  |
| 43 | 60   | 13.3 | 503  | 2 A83027 | hypothetical prote |
| 44 | 59.5 | 13.2 | 256  | 2 G96692 | hypothetical prote |
| 45 | 59.5 | 13.2 | 304  | 2 A83309 | hypothetical prote |

ALIGNMENTS

RESULT 1

I51064  
somatostatin II precursor - rainbow trout  
C;Species: Oncorhynchus mykiss (rainbow trout)  
C;Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 16-Jul-1999  
C;Accession: I51064  
R;Moore, C.A.; Kittilson, J.D.; Dahl, S.K.; Sheridan, M.A.  
Gen. Comp. Endocrinol. 98, 253-261, 1995  
A;Title: Isolation and characterization of a cDNA encoding for preprosomatostatin c  
A;Reference number: I51064; MUID:95354921; PMID:7628684  
A;Accession: I51064  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-115 <MOO>  
A;Cross-references: EMBL:U32471; NID:g975344; PIDN:AAC59695.1; PID:g975345  
C;Superfamily: somatostatin

Query Match 77.2%; Score 347.5; DB 2; Length 115;  
Best Local Similarity 85.7%; Pred. No. 9.6e-32;  
Matches 72; Conservative 2; Mismatches 7; Indels 3; Gaps 2;

QY 1 MRVSIQHCALALLGLALAICSGAASQPDLDLASRRLLQRLAALPHRSGVSRWRTFY 60  
DB 1 MKVCRIHCALALLGLALAICSGAASQPDLDLASRRLLQRLAALPHRSGVSRWRTFY 60  
|||||

QY 61 PNCPCLRWRPRKVKGPQLKAKEDL 84  
DB 61 PNCPCLRWRPRKVKGPQLKAKEDL 81  
|||||

RESULT 2

S20630  
somatostatin precursor - chicken  
C;Species: Gallus gallus (chicken)  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999.#text\_change 10-Sep-1999  
C;Accession: S20630  
R;Nata, K.; Kobayashi, T.; Karahashi, K.; Kato, S.; Yamamoto, H.; Yonekura, H.; Oka  
submitted to the EMBL Data Library, June 1991  
A;Description: Nucleotide sequence determination of chicken somatostatin precursor  
A;Reference number: S20630  
A;Accession: S20630  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-116 <NAT>  
A;Cross-references: EMBL:X60191; NID:g62985; PIDN:CAA42747.1; PID:g62986  
C;Superfamily: somatostatin

Query Match 18.6%; Score 83.5; DB 1; Length 116;  
Best Local Similarity 37.3%; Pred. No. 0.037;  
Matches 31; Conservative 14; Mismatches 25; Indels 13; Gaps 5;

QY 5 QIHCALALLGLALAICSGAA-SQPDLDLASRRLLQRLAALPHRSGVSRWRTFY 63











F:1-24/Domain: signal sequence #status predicted <SIG>  
F:25-88/Domain: propeptide #status predicted <PRO>  
F:89-116/Product: somatostatin-28 #status predicted <M28>  
F:103-116/Product: somatostatin-14 #status predicted <M14>  
F:105-116/Disulfide bonds: #status predicted

Query Match 15.2%; Score 68.5; DB 1; Length 116;  
Best Local Similarity 45.2%; Pred. No. 1.8;  
Matches 19; Conservative 7; Mismatches 11; Indels 5; Gaps 2;

QY 5 QIHCALALLGLALAICS-QGAASQPDLDLASRRLLQALAAA 45  
Db 5 RLQCALAALSIVLALGGVTGAPSDPRL-----RQFLQKSLAAA 42

RESULT 10

A82554  
conserved hypothetical protein XF2472 [imported] - Xylella fastidiosa (strain 9a5c)  
C:Species: Xylella fastidiosa  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Sep-2000  
C:Accession: A82554  
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing  
Nature 406, 151-157, 2000  
Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
Reference number: A82515; MUID:20365717; PMID:10910347  
A:Note: for a complete list of authors see reference number A59328 below  
A:Accession: A82554  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-435 <SIM>  
A:Cross-references: GB:AE004055; GB:AE003849; NID:g9107661; PIDN:AAF85270.1; GSPDB:GN001  
A:Experimental source: strain 9a5c  
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H. as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.  
submitted to GenBank, June 2000  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E. A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A. Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: XF2472  
C:Superfamily: Aquifex aeolicus N-ethylammine chlorohydrolase

Query Match 15.2%; Score 68.5; DB 2; Length 435;  
Best Local Similarity 33.9%; Pred. No. 6.3;  
Matches 20; Conservative 6; Mismatches 18; Indels 15; Gaps 1;

QY 3 VSQIHC-----ALALLGLALAICSQGAASQPDLDLASRRLLQALAAA 46  
Db 267 VSVVHCPESNLKLASGFCFAPALHRAGVNLAIGTDGCASNNDMLSENRTAAMLAKAV 325

RESULT 11

T51335  
subtilisin-like proteinase AIR3, auxin-induced [imported] - Arabidopsis thaliana (fragme  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 03-Nov-2000  
C:Accession: T51335  
R:Neuteboom, L.W.; Ng, J.M.Y.; Kuyper, M.; Clifdesdale, O.R.; Hooykaas, P.J.J.; van der Plant Mol. Biol. 39, 273-287, 1999  
A:Title: Isolation and characterization of cDNA clones corresponding with mRNAs that acc A:Reference number: 225377; MUID:99178779; PMID:10080694  
A:Accession: T51335  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-758 <NEU>

A:Cross-references: EMBL:AF055848; PIDN:AAC62611.1  
A:Experimental source: root culture  
A:Note: accumulates during auxin-induced lateral root formation  
C:Genetics:  
A:Gene: AIR3  
C:Superfamily: subtilisin-like proteinase agl2; subtilisin homology

Query Match 15.1%; Score 68; DB 2; Length 758;  
Best Local Similarity 28.1%; Pred. No. 12;  
Matches 18; Conservative 13; Mismatches 25; Indels 8; Gaps 2;

QY 14 GLALAICSQGAASQPDLDLASRRLLQALAAA--ALPHRSGVSEWRWTFYPCPLRWRP 70  
Db 515 GVSVIAAYTGAVSPTNEQFDRLLFNALSGTSMSCPHISGIAGLLKTRYPs-----WSP 569

QY 71 RKVK 74

Db 570 AAIR 573

RESULT 12

F83244  
hypothetical protein PA3203 [imported] - Pseudomonas aeruginosa (strain PAO1)  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: F83244  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L. ; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pa A:Reference number: A82950; MUID:20437337; PMID:10984043  
A:Accession: F83244  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-120 <STO>  
A:Cross-references: GB:AE004744; GB:AE004091; NID:g9949317; PIDN:AAG06591.1; GSPDB:GN A:Experimental source: strain PAO1  
C:Genetics:  
A:Gene: PA3203

Query Match 14.9%; Score 67; DB 2; Length 120;  
Best Local Similarity 33.9%; Pred. No. 2.7;  
Matches 21; Conservative 8; Mismatches 27; Indels 6; Gaps 2;

QY 1 MRVSIHCALALLGLALAICSQGA-----ASQPDLDLASRRLLQALAAAALPHRSGVSE 55  
Db 2 MRPSAAHCLLSLGLASLAAEAPAPAPIAAQPSITQATAE-LERRLVETERQDELVSR 60

QY 56 WR 57

Db 61 MR 62

RESULT 13

A38943  
nitric-oxide synthase (EC 1.14.13.39), endothelial - bovine  
N:Alternate names: ECNOS; nitric-oxide synthase type III  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 02-Jun-1995 #sequence\_revision 02-Jun-1995 #text\_change 03-Mar-2000  
C:Accession: A38943; A46033; I45945; A42841; I45946; A38944  
R:Lamas, S.; Marsden, P.A.; Li, G.K.; Tempst, P.; Michel, T.  
submitted to GenBank, July 1992  
A:Reference number: A38943  
A:Accession: A38943  
A:Molecule type: mRNA  
A:Residues: 1-1205 <LAM1>  
A:Cross-references: GB:M89952; NID:gl62976; PIDN:AAA30494.1; PID:gl62977  
A:Experimental source: aortic endothelial cells  
R:Lamas, S.; Marsden, P.A.; Li, G.K.; Tempst, P.; Michel, T.  
Proc. Natl. Acad. Sci. U.S.A. 89, 6348-6352, 1992  
A:Title: Endothelial nitric oxide synthase: molecular cloning and characterization of A:Reference number: A46033; MUID:92335295; PMID:1378626



A;Accession: A46033  
A;Status: nucleic acid sequence not shown  
A;Molecule type: mRNA  
A;Residues: 1-860; I', 862-1205 <LAM2>  
A;Cross-references: GB:M89952; NID:g162976  
A;Experimental source: endothelial  
A;Note: sequence extracted from NCBI backbone (NCBIP:108720)  
R;Nishida, K.; Harrison, D.G.; Navas, J.P.; Fisher, A.A.; Dockery, S.P.; Nerem, R.M.; Al  
J. Clin. Invest. 90, 2092-2096, 1992  
A;Title: Molecular Cloning and Characterization of the constitutive bovine aortic Endoth  
A;Reference number: I45945; MUID:93055452; PMID:1385480  
A;Accession: I45945  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-1205 <NIS>  
A;Cross-references: GB:M99057; NID:g163421; PIDN:AAA306657.1; PID:g163422  
R;Sessa, W.C.; Harrison, J.K.; Barber, C.M.; Zeng, D.; Durieux, M.E.; D'Angelo, D.D.; Ly  
J. Biol. Chem. 267, 15274-15276, 1992  
A;Title: Molecular cloning and expression of a cDNA encoding endothelial cell nitric oxi  
A;Reference number: A42841; MUID:92348367; PMID:1379225  
A;Accession: A42841  
A;Status: nucleic acid sequence not shown  
A;Molecule type: mRNA  
A;Residues: 1-51; N', 53-99, 'R', 101-147, 'M', 149-164, 'I', 166-317, 'GA', 320; 'HTGVVRGP', 329-3  
'K', 517-692, 'G', 694-740, 'A', 742-753, 'N', 755-799, 'N', 801-803, 'SA', 806-856, 'V', 858-906, 'LV  
A;Experimental source: aortic endothelial cells  
A;Note: sequence extracted from NCBI backbone (NCBIP:109564); contains a number of typog  
A;Accession: I45946  
A;Status: nucleic acid sequence not shown  
A;Molecule type: mRNA  
A;Residues: 1-99, 'R', 101-164, 'I', 166-317, 'GA', 320, 'HTGVVRGP', 329-454, 'Y', 456-458, 'P', 460  
A;Cross-references: GB:M95674; NID:g163426; PIDN:AAA30669.1; PID:g163427  
A;Experimental source: aortic endothelial cells  
A;Note: submitted to GenBank, August 1992  
A;Note: GenBank entry BOVNOS, release 103.0, has a typographical error in the reference  
C;Function:  
A;Description: catalyzes the oxidation of an L-arginine guanidino nitrogen and of NADPH  
C;Superfamily: nitric-oxide synthase; flavodoxin homology; NADPH-ferrihemoprotein reduct  
C;Keywords: blocked amino end; calmodulin binding; chromoprotein; FAD; flavoprotein; FMN  
F:493-512/Region: calmodulin binding #status predicted  
F:522-1161/Domain: NADPH-ferrihemoprotein reductase homology <FEH>  
F:524-705/Domain: flavodoxin homology <FLX>  
F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
F:3/Modified site: aspartic acid (Asn) #status predicted  
F:186/Binding site: heme iron (Cys) (axial ligand) #status predicted  
  
Query Match 14.9%; Score 67; DB 1; Length 1205;  
Best Local Similarity 35.5%; Pred. No. 25;  
Matches 27; Conservative 4; Mismatches 25; Indels 20; Gaps 5;  
  
QY 8 CALALLGLALAIC-SQGAAS---OPDLASRRLLQRALAAALPHRSGVSRWRTFYPNC 63  
| | | | | : | | | | : | | | | |  
Db 15 CGLG-LGLGLGCGKQGPASPAPEP-----SRAPAPATPHAPDHSPA-----PNS 58  
| | | | | : | | | | : | | | | |  
  
QY 64 PCLWRPRKVKGPOLK 79  
| | | | | : | | | | : | | | | |  
Db 59 PTLTRPPEGPKFPRVK 74  
| | | | | : | | | | : | | | | |  
  
RESULT 14  
H81009  
exodeoxyribonuclease NMB2082 [Imported] - Neisseria meningitidis (strain MC58 serogroup  
C;Species: Neisseria meningitidis  
C;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 02-Feb-2001  
C;Accession: H81009; A82031  
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A  
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;  
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.  
Science 287, 1809-1815, 2000  
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve  
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
A;Reference number: A81000; MUID:20175755; PMID:10710307  
A;Accession: H81009

A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-259 <TET>  
A;Cross-references: GB:AE002557; GB:AE002098; NID:g7227332; PIDN:AAF42400.1; PID:g7  
A;Experimental source: serogroup B, strain MC58  
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.;  
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajanc  
Nature 404, 502-506, 2000  
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 224  
A;Reference number: A81775; MUID:20222556; PMID:10761919  
A;Accession: A82031  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-259 <PAR>  
A;Cross-references: GB:AL162752; GB:AL157959; NID:g7378778; PIDN:CAB83652.1; PID:g7  
A;Experimental source: serogroup A, strain Z2491  
C;Genetics:  
A;Gene: NMB2082; NMA0348  
C;Superfamily: exodeoxyribonuclease III  
  
Query Match 14.8%; Score 66.5; DB 2; Length 259;  
Best Local Similarity 24.6%; Pred. No. 6.4;  
Matches 16; Conservative 13; Mismatches 25; Indels 11; Gaps 2;  
  
QY 14 GLALAICSQGAASQPDLDLASRRLLQRALAAALP-----HRSGVSRWRTFYPNC 63  
| | | | | : | | | | : | | | | |  
Db 141 GRDIVVCGDNIAHQNIDLNKWKGNQKN-SGFLPEREWIGKVIHKLGTDMWRTLYPDV 199  
| | | | | : | | | | : | | | | |  
  
QY 64 PCLRW 68  
| | | | | : | | | | : | | | | |  
Db 200 PGYTW 204  
| | | | | : | | | | : | | | | |  
  
RESULT 15  
T43456  
hypothetical protein DKFZp434L061.1 - human  
C;Species: Homo sapiens (man)  
C;Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 17-Mar-2000  
C;Accession: T43456  
R;Poustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, December 1999  
A;Reference number: Z22516  
A;Accession: T43456  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-571 <AAA>  
A;Cross-references: EMBL:AL133647  
A;Experimental source: adult testis; clone DKFZp434L061  
C;Genetics:  
A;Note: DKFZp434L061.1  
C;Superfamily: hydroxyproline-rich glycoprotein  
  
Query Match 14.8%; Score 66.5; DB 2; Length 571;  
Best Local Similarity 29.0%; Pred. No. 14;  
Matches 20; Conservative 15; Mismatches 19; Indels 15; Gaps 4;  
  
QY 1 MRVSIHCALALLGLALAICSQGAASQPDLDLASRRLLQRALAAA----LPHRSGVSRW 56  
| | | | | : | | | | : | | | | |  
Db 405 LQLSEVN-----IPLSLGVCPLG-----PVPLTKELYQQAAMEAAWHMHPD-SERI 453  
| | | | | : | | | | : | | | | |  
  
QY 57 RTFYPNCPC 65  
| | | | | : | | | | : | | | | |  
Db 454 RQYLPRNPC 462  
| | | | | : | | | | : | | | | |  
  
Search completed: March 21, 2003, 11:41:40  
Job time : 20.4549 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2003, 11:53:33 ; Search time 16.6094 Seconds  
(without alignments)  
276.816 Million cell updates/sec

Title: US-09-727-739B-17  
Perfect score: 450  
Sequence: 1 MRVSIQHICALALLGLAIC.....RWRPRKVGPKLAKEDLER 86

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 221153 seqs, 53462247 residues

Total number of hits satisfying chosen parameters: 221153

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID                     | Description        |
|------------|-------|-------------|--------|------------------------|--------------------|
| 1          | 67.5  | 15.0        | 95     | 10 US-09-864-761-41361 | Sequence 41361, A  |
| 2          | 65.5  | 14.6        | 513    | 10 US-09-833-745-63    | Sequence 63, Appl  |
| 3          | 63.5  | 14.1        | 333    | 9 US-09-984-271-131    | Sequence 131, Appl |
| 4          | 62.5  | 13.9        | 110    | 10 US-09-766-396-3     | Sequence 3, Appl   |
| 5          | 62.5  | 13.9        | 110    | 12 US-10-062-375-3     | Sequence 3, Appl   |
| 6          | 62    | 13.8        | 833    | 9 US-09-470-276-54     | Sequence 54, Appl  |
| 7          | 61.5  | 13.7        | 333    | 9 US-10-001-054-42     | Sequence 42, Appl  |
| 8          | 61.5  | 13.7        | 333    | 9 US-10-028-072-132    | Sequence 132, Appl |
| 9          | 61.5  | 13.7        | 333    | 9 US-10-121-049-132    | Sequence 132, Appl |
| 10         | 61.5  | 13.7        | 333    | 9 US-10-123-904-132    | Sequence 132, Appl |
| 11         | 61.5  | 13.7        | 333    | 9 US-10-140-470-132    | Sequence 132, Appl |
| 12         | 61.5  | 13.7        | 333    | 9 US-10-175-746-132    | Sequence 132, Appl |
| 13         | 61.5  | 13.7        | 333    | 9 US-10-176-918-132    | Sequence 132, Appl |
| 14         | 61.5  | 13.7        | 333    | 9 US-10-176-921-132    | Sequence 132, Appl |
| 15         | 61.5  | 13.7        | 333    | 9 US-10-137-865-132    | Sequence 132, Appl |
| 16         | 61.5  | 13.7        | 333    | 9 US-10-140-474-132    | Sequence 132, Appl |
| 17         | 61.5  | 13.7        | 333    | 9 US-10-142-431-132    | Sequence 132, Appl |
| 18         | 61.5  | 13.7        | 333    | 9 US-10-143-114-132    | Sequence 132, Appl |
| 19         | 61.5  | 13.7        | 333    | 9 US-10-140-002-132    | Sequence 132, Appl |

|    |      |      |      |    |                    |                   |
|----|------|------|------|----|--------------------|-------------------|
| 20 | 61.5 | 13.7 | 333  | 9  | US-10-142-419-132  | Sequence 132, App |
| 21 | 61.5 | 13.7 | 333  | 9  | US-10-123-262-132  | Sequence 132, App |
| 22 | 61.5 | 13.7 | 333  | 9  | US-10-142-423-132  | Sequence 132, App |
| 23 | 61   | 13.6 | 697  | 10 | US-09-770-643A-18  | Sequence 18, Appl |
| 24 | 61   | 13.6 | 745  | 10 | US-09-770-643A-16  | Sequence 16, Appl |
| 25 | 61   | 13.6 | 791  | 10 | US-09-770-643A-22  | Sequence 22, Appl |
| 26 | 61   | 13.6 | 839  | 10 | US-09-770-643A-20  | Sequence 20, Appl |
| 27 | 61   | 13.6 | 1175 | 10 | US-09-770-643A-26  | Sequence 26, Appl |
| 28 | 61   | 13.6 | 1175 | 10 | US-09-770-643A-30  | Sequence 30, Appl |
| 29 | 61   | 13.6 | 1259 | 10 | US-09-770-643A-4   | Sequence 4, Appl  |
| 30 | 61   | 13.6 | 1298 | 10 | US-09-770-643A-24  | Sequence 24, Appl |
| 31 | 61   | 13.6 | 1298 | 10 | US-09-770-643A-28  | Sequence 28, Appl |
| 32 | 61   | 13.6 | 1307 | 10 | US-09-770-643A-2   | Sequence 2, Appl  |
| 33 | 60.5 | 13.4 | 90   | 10 | US-09-789-561-128  | Sequence 128, App |
| 34 | 59   | 13.1 | 468  | 9  | US-09-764-868-665  | Sequence 665, App |
| 35 | 58.5 | 13.0 | 396  | 8  | US-08-900-220C-17  | Sequence 17, Appl |
| 36 | 58.5 | 13.0 | 396  | 9  | US-09-883-848A-17  | Sequence 17, Appl |
| 37 | 58.5 | 13.0 | 396  | 10 | US-09-244-466-2    | Sequence 2, Appl  |
| 38 | 58.5 | 13.0 | 396  | 10 | US-09-151-999-17   | Sequence 17, Appl |
| 39 | 58   | 12.9 | 234  | 9  | US-09-934-392-6    | Sequence 6, Appl  |
| 40 | 58   | 12.9 | 274  | 10 | US-09-739-907-171  | Sequence 171, App |
| 41 | 58   | 12.9 | 525  | 10 | US-09-739-907-79   | Sequence 79, Appl |
| 42 | 58   | 12.9 | 1503 | 9  | US-10-007-706-1    | Sequence 1, Appl  |
| 43 | 57.5 | 12.8 | 384  | 10 | US-09-815-242-5184 | Sequence 5184, Ap |
| 44 | 57   | 12.7 | 834  | 9  | US-09-470-276-2    | Sequence 2, Appl  |
| 45 | 57   | 12.7 | 834  | 10 | US-09-934-909-29   | Sequence 29, Appl |

ALIGNMENTS

RESULT 1  
US-09-864-761-41361  
; Sequence 41361, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL  
; FILE REFERENCE: Acomica-x-1  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US/09/864,761  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661







REFERENCE/DOCKET NUMBER: 22908-0002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-7041  
TELEFAX: (415) 324-0638  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 110 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: C-terminal  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-09-766-396-3

Query Match 13.9%; Score 62.5; DB 10; Length 110;  
Best Local Similarity 48.7%; Pred. No. 3.7;  
Matches 19; Conservative 5; Mismatches 10; Indels 5; Gaps 2;

QY 8 CALALLGLALAICS-QGAASQPDLDLASRRLLQALAAA 45  
Db 2 CALAALCIVLALGGVTGAPSDPRL---RQFLQKSLAAA 36

RESULT 5

US-10-062-375-3  
Sequence 3, Application US/10062375  
Patent No. US20020133000A1

GENERAL INFORMATION:

APPLICANT: Sutcliffe, Gregor J.

de Lecea, Luis

Siggins, George R.

Henriksen, Steven J.

TITLE OF INVENTION: CORRELATION: NEUROPEPTIDES,

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE

STREET: 10666 NO. US20020133000A1th Torrey Pines Road, TPC-8

CITY: La Jolla

STATE: California

COUNTRY: US

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/062,375

FILING DATE: 30-Jan-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION NUMBER: US/08/857,389

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Schmonsees, William

REGISTRATION NUMBER: 31,796

REFERENCE/DOCKET NUMBER: 22908-0002

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-7041

TELEFAX: (415) 324-0638

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 110 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

FRAGMENT TYPE: C-terminal

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-10-062-375-3

Query Match 13.9%; Score 62.5; DB 12; Length 110;  
Best Local Similarity 48.7%; Pred. No. 3.7;

Matches 19; Conservative 5; Mismatches 10; Indels 5; Gaps 2;  
QY 8 CALALLGLALAICS-QGAASQPDLDLASRRLLQALAAA 45  
Db 2 CALAALCIVLALGGVTGAPSDPRL---RQFLQKSLAAA 36

RESULT 6

US-09-470-276-54

Sequence 54, Application US/09470276

Publication No. US20020187469A1

GENERAL INFORMATION:

APPLICANT: DANA-FARBER CANCER INSTITUTE, INC.

APPLICANT: KOLODNER, Richard

APPLICANT: WINAND, Nena

TITLE OF INVENTION: A METHOD OF DETECTION OF ALTERATIONS IN MSH5

FILE REFERENCE: 700157/47483C

CURRENT APPLICATION NUMBER: US/09/470,276

PRIOR FILING DATE: 1999-12-22

PRIOR APPLICATION NUMBER: 60/051,686

PRIOR FILING DATE: 1997-07-03

PRIOR APPLICATION NUMBER: PCT/US98/13850

PRIOR FILING DATE: 1998-07-02

NUMBER OF SEQ ID NOS: 104

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 54

LENGTH: 833

TYPE: PRT

ORGANISM: Human

US-09-470-276-54

Query Match 13.8%; Score 62; DB 9; Length 833;

Best Local Similarity 27.5%; Pred. No. 42;

Matches 30; Conservative 12; Mismatches 37; Indels 30; Gaps 5;

QY 3 VSGIHC-----ALALLGLALAICSQGAASQPDLDLASRRLLQALAAAAL-----PHR 49

Db 489 LGDLHCEIRDQETLLMYQLQCQVLARASVLTVDLASRLDVLALASAARDYGYSRPHY 548

QY 50 S-----GVSER-----WRTFYPN---CPCLRWRPRKVKGPQLKAK 81

Db 549 SPCHGVRIIRNGRHPMLMELCARTFVPNSTDCGGDQGRVKVITGPNSSGK 597

RESULT 7

US-10-001-054-42

Sequence 42, Application US/10001054

Publication No. US20020192209A1

GENERAL INFORMATION:

APPLICANT: Genentech, Inc.

APPLICANT: Baker, Kevin

APPLICANT: Goddard, Audrey

APPLICANT: Gurney, Austin

APPLICANT: Hebert, Carolyn

APPLICANT: Hensel, William

APPLICANT: Kabakoff, Rhona

APPLICANT: Shelton, David

APPLICANT: Smith, Victoria

APPLICANT: Watanabe, Colin

APPLICANT: Wood, William

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING NEOPLASTIC

TITLE OF INVENTION: CELL GROWTH

FILE REFERENCE: P3034R1PCT

CURRENT APPLICATION NUMBER: US/10/001,054

CURRENT FILING DATE: 2001-11-30

PRIOR APPLICATION NUMBER: 60/059114

PRIOR FILING DATE: 1997-09-17

PRIOR APPLICATION NUMBER: 60/079689

PRIOR FILING DATE: 1998-03-27

PRIOR APPLICATION NUMBER: 60/079920

PRIOR FILING DATE: 1998-03-30

PRIOR APPLICATION NUMBER: 60/082999

PRIOR FILING DATE: 1998-04-24



; PRIOR APPLICATION NUMBER: 60/083545  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/085149  
; PRIOR FILING DATE: 1998-05-12  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/088858  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/090691  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/096891  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: 60/096894  
; PRIOR FILING DATE: 1998-08-17  
; PRIOR APPLICATION NUMBER: 60/099803  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/100263  
; PRIOR FILING DATE: 1998-09-14  
; PRIOR APPLICATION NUMBER: 60/100390  
; PRIOR FILING DATE: 1998-09-15  
; PRIOR APPLICATION NUMBER: 60/101476  
; PRIOR FILING DATE: 1998-09-23  
; PRIOR APPLICATION NUMBER: 60/107783  
; PRIOR FILING DATE: 1998-11-10  
; PRIOR APPLICATION NUMBER: 60/108849  
; PRIOR FILING DATE: 1998-11-18  
; PRIOR APPLICATION NUMBER: 60/112420  
; PRIOR FILING DATE: 1998-12-15  
; PRIOR APPLICATION NUMBER: 60/113296  
; PRIOR FILING DATE: 1998-12-22  
; PRIOR APPLICATION NUMBER: 60/115554  
; PRIOR FILING DATE: 1999-01-12  
; PRIOR APPLICATION NUMBER: 60/115558  
; PRIOR FILING DATE: 1999-01-12  
; PRIOR APPLICATION NUMBER: 60/116533  
; PRIOR FILING DATE: 1999-01-20  
; PRIOR APPLICATION NUMBER: 60/123618  
; PRIOR FILING DATE: 1999-03-10  
; PRIOR APPLICATION NUMBER: 60/131294  
; PRIOR FILING DATE: 1999-04-07  
; PRIOR APPLICATION NUMBER: 60/140650  
; PRIOR FILING DATE: 1999-06-22  
; PRIOR APPLICATION NUMBER: 60/141037  
; PRIOR FILING DATE: 1999-06-23  
; PRIOR APPLICATION NUMBER: 60/144758  
; PRIOR FILING DATE: 1999-07-20  
; PRIOR APPLICATION NUMBER: 60/162506  
; PRIOR FILING DATE: 1999-10-29  
; PRIOR APPLICATION NUMBER: 60/170262  
; PRIOR FILING DATE: 1999-12-09  
; PRIOR APPLICATION NUMBER: 60/187202  
; PRIOR FILING DATE: 2000-03-03  
; PRIOR APPLICATION NUMBER: 60/209832  
; PRIOR FILING DATE: 2000-06-05  
; PRIOR APPLICATION NUMBER: 60/232887  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: 09/180997  
; PRIOR FILING DATE: 1998-11-19  
; PRIOR APPLICATION NUMBER: 09/218517  
; PRIOR FILING DATE: 1998-12-22  
; PRIOR APPLICATION NUMBER: 09/284291  
; PRIOR FILING DATE: 1999-04-12  
; PRIOR APPLICATION NUMBER: 09/380137  
; PRIOR FILING DATE: 1999-08-25  
; PRIOR APPLICATION NUMBER: 09/380138  
; PRIOR FILING DATE: 1999-08-25  
; PRIOR APPLICATION NUMBER: 09/380913  
; PRIOR FILING DATE: 1999-09-09  
; PRIOR APPLICATION NUMBER: 09/403297  
; PRIOR FILING DATE: 1999-10-18  
; PRIOR APPLICATION NUMBER: 09/423741  
; PRIOR FILING DATE: 1999-11-10  
; PRIOR APPLICATION NUMBER: 09/709238

; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 09/802706  
; PRIOR FILING DATE: 2001-03-09  
; PRIOR APPLICATION NUMBER: 09/866034  
; PRIOR FILING DATE: 2001-05-25  
; PRIOR APPLICATION NUMBER: 09/872035  
; PRIOR FILING DATE: 2001-06-01  
; PRIOR APPLICATION NUMBER: 09/882636  
; PRIOR FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: 09/918585  
; PRIOR FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: 09/924419  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: 09/927796  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: 09/929404  
; PRIOR FILING DATE: 2001-08-13  
; PRIOR APPLICATION NUMBER: 09/941992  
; PRIOR FILING DATE: 2001-08-28  
; PRIOR APPLICATION NUMBER: 09/946374  
; PRIOR FILING DATE: 2001-09-04  
; PRIOR APPLICATION NUMBER: PCT/US98/18824  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: PCT/US99/00106  
; PRIOR FILING DATE: 1999-01-05  
; PRIOR APPLICATION NUMBER: PCT/US99/05028  
; PRIOR FILING DATE: 1999-03-08  
; PRIOR APPLICATION NUMBER: PCT/US99/08615  
; PRIOR FILING DATE: 1999-04-20  
; PRIOR APPLICATION NUMBER: PCT/US99/12252  
; PRIOR FILING DATE: 1999-06-02  
; PRIOR APPLICATION NUMBER: PCT/US99/20111  
; PRIOR FILING DATE: 1999-09-01  
; PRIOR APPLICATION NUMBER: PCT/US99/20594  
; PRIOR FILING DATE: 1999-09-08  
; PRIOR APPLICATION NUMBER: PCT/US99/28313  
; PRIOR FILING DATE: 1999-11-30  
; PRIOR APPLICATION NUMBER: PCT/US99/28551  
; PRIOR FILING DATE: 1999-12-02  
; PRIOR APPLICATION NUMBER: PCT/US99/28634  
; PRIOR FILING DATE: 1999-12-01  
; PRIOR APPLICATION NUMBER: PCT/US99/30095  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: PCT/US99/30999  
; PRIOR FILING DATE: 1999-12-20  
; PRIOR APPLICATION NUMBER: PCT/US00/00376  
; PRIOR FILING DATE: 2000-01-06  
; PRIOR APPLICATION NUMBER: PCT/US00/03565  
; PRIOR FILING DATE: 2000-02-11  
; PRIOR APPLICATION NUMBER: PCT/US00/04341  
; PRIOR FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: PCT/US00/04342  
; PRIOR FILING DATE: 2000-03-02  
; PRIOR APPLICATION NUMBER: PCT/US00/06884  
; PRIOR FILING DATE: 2000-03-15  
; PRIOR APPLICATION NUMBER: PCT/US00/08439  
; PRIOR FILING DATE: 2000-03-30  
; PRIOR APPLICATION NUMBER: PCT/US00/13705  
; PRIOR FILING DATE: 2000-05-17  
; PRIOR APPLICATION NUMBER: PCT/US00/14042  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: PCT/US00/14941  
; PRIOR FILING DATE: 2000-05-30  
; PRIOR APPLICATION NUMBER: PCT/US00/15264  
; PRIOR FILING DATE: 2000-06-02  
; PRIOR APPLICATION NUMBER: PCT/US00/22031  
; PRIOR FILING DATE: 2000-08-11  
; PRIOR APPLICATION NUMBER: PCT/US00/23522  
; PRIOR FILING DATE: 2000-08-23  
; PRIOR APPLICATION NUMBER: PCT/US00/30873  
; PRIOR FILING DATE: 2000-11-10



















GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2003, 11:41:02 ; Search time 16.2403 Seconds  
(without alignments)  
155.808 Million cell updates/sec

Title: US-09-727-739B-17  
Perfect score: 450  
Sequence: 1 MRVSIHICALALLGLALAIC.....RWRPRKVKGPQLKAKEDLER 86

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents.AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 67    | 14.9        | 1205   | 1     | US-07-908-245-2    |
| 2          | 67    | 14.9        | 1205   | 2     | US-08-319-866-10   |
| 3          | 67    | 14.9        | 1205   | 4     | US-09-123-708-6    |
| 4          | 67    | 14.9        | 1205   | 4     | US-09-123-624-6    |
| 5          | 62.5  | 13.9        | 110    | 3     | US-08-648-322-3    |
| 6          | 59.5  | 13.2        | 2476   | 2     | US-08-276-967-2    |
| 7          | 58.5  | 13.0        | 396    | 4     | US-09-325-256-24   |
| 8          | 58    | 12.9        | 802    | 4     | US-09-632-098-2    |
| 9          | 57    | 12.7        | 479    | 1     | US-08-484-105-10   |
| 10         | 57    | 12.7        | 479    | 1     | US-08-484-106-10   |
| 11         | 57    | 12.7        | 834    | 4     | US-09-143-571-29   |
| 12         | 56.5  | 12.6        | 47     | 4     | US-08-469-260A-344 |
| 13         | 56.5  | 12.6        | 396    | 1     | US-08-176-427B-4   |
| 14         | 56.5  | 12.6        | 396    | 2     | US-08-356-060A-9   |
| 15         | 56.5  | 12.6        | 396    | 4     | US-08-460-900C-9   |
| 16         | 56.5  | 12.6        | 396    | 4     | US-08-674-509B-9   |
| 17         | 56.5  | 12.6        | 396    | 4     | US-08-954-698-9    |
| 18         | 56.5  | 12.6        | 396    | 4     | US-09-293-505-13   |
| 19         | 56.5  | 12.6        | 396    | 4     | US-08-957-874-9    |
| 20         | 56.5  | 12.6        | 396    | 4     | US-09-325-256-18   |
| 21         | 55.5  | 12.3        | 1296   | 3     | US-08-728-603-15   |
| 22         | 54.5  | 12.1        | 741    | 2     | US-08-462-481-2    |
| 23         | 54.5  | 12.1        | 741    | 2     | US-08-436-771-2    |
| 24         | 54.5  | 12.1        | 741    | 2     | US-08-436-771-4    |
| 25         | 54.5  | 12.1        | 741    | 2     | US-08-434-998-2    |
| 26         | 54.5  | 12.1        | 741    | 2     | US-08-434-998-4    |
| 27         | 54.5  | 12.1        | 741    | 2     | US-08-487-797-2    |

|    |      |      |     |   |                   |                    |
|----|------|------|-----|---|-------------------|--------------------|
| 28 | 54.5 | 12.1 | 741 | 2 | US-08-487-797-4   | Sequence 4, Appli  |
| 29 | 54.5 | 12.1 | 741 | 2 | US-08-701-005A-2  | Sequence 2, Appli  |
| 30 | 54.5 | 12.1 | 741 | 2 | US-08-479-895-2   | Sequence 2, Appli  |
| 31 | 54.5 | 12.1 | 741 | 3 | US-08-943-956A-2  | Sequence 2, Appli  |
| 32 | 54.5 | 12.1 | 741 | 5 | PCT-US95-02058-2  | Sequence 2, Appli  |
| 33 | 54.5 | 12.1 | 741 | 5 | PCT-US95-02058-4  | Sequence 4, Appli  |
| 34 | 54   | 12.0 | 315 | 4 | US-09-370-398-6   | Sequence 6, Appli  |
| 35 | 54   | 12.0 | 374 | 1 | US-08-468-847B-12 | Sequence 12, Appli |
| 36 | 54   | 12.0 | 375 | 2 | US-08-459-101A-2  | Sequence 2, Appli  |
| 37 | 54   | 12.0 | 408 | 3 | US-08-458-434A-2  | Sequence 2, Appli  |
| 38 | 54   | 12.0 | 671 | 4 | US-09-132-118-2   | Sequence 2, Appli  |
| 39 | 54   | 12.0 | 709 | 1 | US-08-444-005-17  | Sequence 17, Appli |
| 40 | 53.5 | 11.9 | 112 | 4 | US-09-344-587-14  | Sequence 14, Appli |
| 41 | 53.5 | 11.9 | 120 | 1 | US-08-347-492B-2  | Sequence 2, Appli  |
| 42 | 53.5 | 11.9 | 120 | 2 | US-08-798-143-2   | Sequence 2, Appli  |
| 43 | 53.5 | 11.9 | 120 | 5 | PCT-US95-15484-2  | Sequence 2, Appli  |
| 44 | 53.5 | 11.9 | 136 | 5 | PCT-US95-07171-2  | Sequence 2, Appli  |
| 45 | 53.5 | 11.9 | 191 | 5 | PCT-US95-07171-3  | Sequence 3, Appli  |

ALIGNMENTS

RESULT 1  
US-07-908-245-2  
; Sequence 2, Application US/07908245  
; Patent No. 5498539  
; GENERAL INFORMATION:  
; APPLICANT: Harrison, David G.  
; APPLICANT: Alexander, R. Wayne  
; APPLICANT: Murphy, T.J.  
; APPLICANT: Nishida, Ken'ichi  
; TITLE OF INVENTION: Endothelial Nitric Oxide Synthase  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Kilpatrick & Cody  
; STREET: 1100 Peachtree Street, Suite 2800  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: U.S.  
; ZIP: 30309-4530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/908,245  
; FILING DATE: 19920702  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pabst, Patrea L.  
; REGISTRATION NUMBER: 31,284  
; REFERENCE/DOCKET NUMBER: EMU 111  
; TELEPHONE: 404-815-6508  
; TELEFAX: 404-815-6555  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1205 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
; ORIGINAL SOURCE:  
; ORGANISM: Bovine  
; TISSUE TYPE: Aorta  
; CELL TYPE: Endothelial  
; FEATURE:  
; NAME/KEY: Binding-site



```
LOCATION: 496..512 /note= "CA++/CAM binding domain"
OTHER INFORMATION:
FEATURE:
NAME/KEY: Binding-site
LOCATION: 651..678 /note= "FMN binding domain"
OTHER INFORMATION:
FEATURE:
NAME/KEY: Binding-site
LOCATION: 795..806 /note= "FAD-Pyrophosphate binding"
OTHER INFORMATION:
FEATURE:
NAME/KEY: Binding-site
LOCATION: 937..947 /note= "FAD-Isolalloxanthine binding domain"
OTHER INFORMATION:
FEATURE:
NAME/KEY: Binding-site
LOCATION: 1012..1030 /note= "NADPH-Ribose binding"
OTHER INFORMATION:
FEATURE:
NAME/KEY: Binding-site
LOCATION: 1111..1124 /note= "NADPH-Ribose binding"
OTHER INFORMATION:
FEATURE:
NAME/KEY: Domain
LOCATION: 33..34 /note= "Potential proline directed phosphorylation site"
OTHER INFORMATION:
FEATURE:
NAME/KEY: Domain
LOCATION: 46..47 /note= "Potential proline directed phosphorylation site"
OTHER INFORMATION:
FEATURE:
NAME/KEY: Domain
LOCATION: 53..54 /note= "Potential proline directed phosphorylation site"
OTHER INFORMATION:
FEATURE:
NAME/KEY: Domain
LOCATION: 58..59 /note= "Potential proline directed phosphorylation site"
OTHER INFORMATION:
FEATURE:
NAME/KEY: Domain
LOCATION: 97..98 /note= "Potential proline directed phosphorylation site"
OTHER INFORMATION:
FEATURE:
NAME/KEY: Domain
LOCATION: 116..117 /note= "Potential proline directed phosphorylation site"
OTHER INFORMATION:
FEATURE:
NAME/KEY: Domain
LOCATION: 282..283 /note= "Potential proline phosphorylation site"
OTHER INFORMATION:
FEATURE:
NAME/KEY: Domain
LOCATION: 459..460 /note= "Potential proline directed phosphorylation site"
OTHER INFORMATION:
FEATURE:
NAME/KEY: Domain
LOCATION: 472..473 /note= "Potential proline directed phosphorylation site"
OTHER INFORMATION:
FEATURE:
NAME/KEY: Domain
```

```
LOCATION: 602..603 /note= "Potential proline directed phosphorylation site"
OTHER INFORMATION:
FEATURE:
NAME/KEY: Domain
LOCATION: 727..728 /note= "Potential proline directed phosphorylation site"
OTHER INFORMATION:
FEATURE:
NAME/KEY: Domain
LOCATION: 838..839 /note= "Potential proline directed phosphorylation site"
OTHER INFORMATION:
FEATURE:
NAME/KEY: Domain
LOCATION: 869..870 /note= "Potential proline directed phosphorylation site"
OTHER INFORMATION:
FEATURE:
NAME/KEY: Domain
LOCATION: 872..873 /note= "Potential proline directed phosphorylation site"
OTHER INFORMATION:
FEATURE:
NAME/KEY: Domain
LOCATION: 1085..1086 /note= "Potential proline directed phosphorylation site"
OTHER INFORMATION:
FEATURE:
NAME/KEY: Domain
LOCATION: 1202..1203 /note= "Potential proline directed phosphorylation site"
OTHER INFORMATION:
FEATURE:
NAME/KEY: Domain
LOCATION: 114..116 /note= "cAMP dependent phosphorylation site"
OTHER INFORMATION:
FEATURE:
NAME/KEY: Domain
LOCATION: 141..143 /note= "cAMP dependent phosphorylation site"
OTHER INFORMATION:
FEATURE:
NAME/KEY: Domain
LOCATION: 168..170 /note= "cAMP dependent phosphorylation site"
OTHER INFORMATION:
FEATURE:
NAME/KEY: Domain
LOCATION: 633..635 /note= "cAMP dependent phosphorylation site"
OTHER INFORMATION:
FEATURE:
NAME/KEY: Domain
LOCATION: 836..838 /note= "cAMP dependent phosphorylation site"
OTHER INFORMATION:
FEATURE:
NAME/KEY: Domain
LOCATION: 1051..1053 /note= "cAMP dependent phosphorylation site"
OTHER INFORMATION:
FEATURE:
NAME/KEY: Domain
LOCATION: 738..740 /note= "cAMP dependent phosphorylation site"
OTHER INFORMATION:
FEATURE:
NAME/KEY: Domain
```

US-07-908-245-2

Query Match 14.9%; Score 67; DB 1; Length 1205;  
Best Local Similarity 35.5%; Pred. No. 7;  
Matches 27; Conservative 4; Mismatches 25; Indels 20; Gaps 5;



QY 8 CALALLGLALAIC-SQGAAS---OPDLDLASRRLLQRALAAALPHRSGVSRWRTFFPNC 63  
| | | | | : | | | | : | | | | |  
Db 15 CGLG-LGLGLGCGKQGPASPAPEP-----SRAPATPHADHSPA-----PNS 58  
QY 64 PCLRWRPRKVKGPOLK 79  
| | | | | : | | | | : | | | | |  
Db 59 PTLTRPPEGPKFPRVK 74

RESULT 2  
US-08-319-866-10  
; Sequence 10, Application US/08319866  
; Patent No. 5929223  
; GENERAL INFORMATION:  
; APPLICANT: Tully, Timothy P.  
; APPLICANT: Yin, Jerry C.  
; APPLICANT: Regulski, Michael  
; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF GENES  
; TITLE OF INVENTION: ASSOCIATED WITH LONG-TERM MEMORY  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
; STREET: Two Militia Drive  
; CITY: Lexington  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02173  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/319,866  
; FILING DATE: 7-OCT-1994  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Granahan, Patricia  
; REGISTRATION NUMBER: 32,227  
; REFERENCE/DOCKET NUMBER: CSHL94-03  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 861-6240  
; TELEFAX: (617) 861-9540  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1205 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-319-866-10

Query Match 14.9%; Score 67; DB 2; Length 1205;  
Best Local Similarity 35.5%; Pred. No. 7;  
Matches 27; Conservative 4; Mismatches 25; Indels 20; Gaps 5;  
QY 8 CALALLGLALAIC-SQGAAS---OPDLDLASRRLLQRALAAALPHRSGVSRWRTFFPNC 63  
| | | | | : | | | | : | | | | |  
Db 15 CGLG-LGLGLGCGKQGPASPAPEP-----SRAPATPHADHSPA-----PNS 58  
QY 64 PCLRWRPRKVKGPOLK 79  
| | | | | : | | | | : | | | | |  
Db 59 PTLTRPPEGPKFPRVK 74

RESULT 3  
US-09-123-708-6  
; Sequence 6, Application US/09123708  
; Patent No. 6146887

; GENERAL INFORMATION:  
; APPLICANT: SCHRAEDER, Juergen  
; APPLICANT: GOECKE, Axel  
; TITLE OF INVENTION: DNA EXPRESSION VECTORS FOR USE IN GENE THERAPEUTIC  
; TITLE OF INVENTION: TREATMENT OF VASCULAR DISORDERS  
; FILE REFERENCE: 511169-2003  
; CURRENT APPLICATION NUMBER: US/09/123,708  
; CURRENT FILING DATE: 1998-07-28  
; EARLIER APPLICATION NUMBER: 08/553,503  
; EARLIER FILING DATE: 1996-03-01  
; EARLIER APPLICATION NUMBER: P4411402.8  
; EARLIER FILING DATE: 1994-03-31  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 1205  
; TYPE: PRT  
; ORGANISM: Cytomegalovirus  
US-09-123-708-6

Query Match 14.9%; Score 67; DB 4; Length 1205;  
Best Local Similarity 35.5%; Pred. No. 7;  
Matches 27; Conservative 4; Mismatches 25; Indels 20; Gaps 5;  
QY 8 CALALLGLALAIC-SQGAAS---OPDLDLASRRLLQRALAAALPHRSGVSRWRTFFPNC 63  
| | | | | : | | | | : | | | | |  
Db 15 CGLG-LGLGLGCGKQGPASPAPEP-----SRAPATPHADHSPA-----PNS 58  
QY 64 PCLRWRPRKVKGPOLK 79  
| | | | | : | | | | : | | | | |  
Db 59 PTLTRPPEGPKFPRVK 74

RESULT 4  
US-09-123-624-6  
; Sequence 6, Application US/09123624  
; Patent No. 6149936  
; GENERAL INFORMATION:  
; APPLICANT: SCHRAEDER, Juergen  
; APPLICANT: GOECKE, Axel  
; TITLE OF INVENTION: DNA EXPRESSION VECTORS FOR USE IN THE GENE THERAPEUTIC  
; TITLE OF INVENTION: TREATMENT OF VASCULAR DISORDERS  
; FILE REFERENCE: 511169-2004  
; CURRENT APPLICATION NUMBER: US/09/123,624  
; CURRENT FILING DATE: 1998-07-28  
; PRIOR APPLICATION NUMBER: 08/553,503  
; PRIOR FILING DATE: 1996-03-01  
; PRIOR APPLICATION NUMBER: 4411402.8  
; PRIOR FILING DATE: 1994-03-31  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 1205  
; TYPE: PRT  
; ORGANISM: Bos taurus  
US-09-123-624-6

Query Match 14.9%; Score 67; DB 4; Length 1205;  
Best Local Similarity 35.5%; Pred. No. 7;  
Matches 27; Conservative 4; Mismatches 25; Indels 20; Gaps 5;  
QY 8 CALALLGLALAIC-SQGAAS---OPDLDLASRRLLQRALAAALPHRSGVSRWRTFFPNC 63  
| | | | | : | | | | : | | | | |  
Db 15 CGLG-LGLGLGCGKQGPASPAPEP-----SRAPATPHADHSPA-----PNS 58  
QY 64 PCLRWRPRKVKGPOLK 79  
| | | | | : | | | | : | | | | |  
Db 59 PTLTRPPEGPKFPRVK 74

RESULT 5  
US-08-648-322-3  
; Sequence 3, Application US/08648322



Patent No. 6074872  
GENERAL INFORMATION:  
APPLICANT: Sutcliffe, Gregor J.  
APPLICANT: de Lecea, Luis  
TITLE OF INVENTION: CORTISTATIN: NEUROPEPTIDES,  
NUMBER OF INVENTION: COMPOSITIONS AND METHODS  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE  
STREET: 10666 No. 6074872th Torrey Pines Road, TPC-8  
CITY: La Jolla  
STATE: California  
COUNTRY: US  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/648,322  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: 519.0  
TELEPHONE: (619) 554-2937  
TELEFAX: (619) 554-6312  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 110 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: C-terminal  
US-08-648-322-3

Query Match 13.9%; Score 62.5; DB 3; Length 110;  
Best Local Similarity 48.7%; Pred. No. 1.2;  
Matches 19; Conservative 5; Mismatches 10; Indels 5; Gaps 2;

QY 8 CALALLGLALAICS-QGAASQPDLDLASRRLLQRALAAA 45  
| | | | | : | | : | | | | : | | : | | | |  
Db 2 CALAALCIVLALGGVTGAPSDPRL----RQFLQKSLAAA 36

RESULT 6  
US-08-276-967-2  
Sequence 2, Application US/08276967  
Patent No. 5851817  
GENERAL INFORMATION:  
APPLICANT: Hardy, Daniel M.  
APPLICANT: Garbers, David L.  
TITLE OF INVENTION: Species-Specific Egg-Binding Proteins of  
TITLE OF INVENTION: Sperm  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P. O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210-4433  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/276,967  
FILING DATE: Submitted Herewith

CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Kitchell, Barbara S.  
REGISTRATION NUMBER: 33,928  
REFERENCE/DOCKET NUMBER: UTSD:418\KIT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 713-787-1400  
TELEFAX: 713-789-2679  
TELEX: 79-0924  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2476 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-276-967-2  
Query Match 13.2%; Score 59.5; DB 2; Length 2476;  
Best Local Similarity 26.2%; Pred. No. 1.6e+02;  
Matches 17; Conservative 7; Mismatches 24; Indels 17; Gaps 3;

QY 20 CSQGAASQPDLDLASRRLLQRALAAALPHRSG---VSEWRWTFYPNC-----P 64  
| | | | | : | | : | | | | : | | : | | | |  
Db 1100 CVEGCECDPGFVLSGLQCVSRSECGCLDSTAGYVKVGERW--FKPGCRQLCICEGNNRTR 1157  
QY 65 CLRWR 69  
| | |  
Db 1158 CVLWR 1162

RESULT 7  
US-09-325-256-24  
Sequence 24, Application US/09325256  
Patent No. 6444793  
GENERAL INFORMATION:  
APPLICANT: PEPINSKY, R. BLAKE  
APPLICANT: BAKER, DARREN P.  
APPLICANT: WEN, DINGYI  
APPLICANT: WILLIAMS, KEVIN P.  
APPLICANT: GARGER, ELLEN A.  
APPLICANT: TAYLOR, FREDERICK R.  
APPLICANT: GALDES, ALPHONSE  
APPLICANT: PORTER, JEFFREY  
TITLE OF INVENTION: HYDROPHOBICALLY-MODIFIED PROTEIN COMPOSITIONS AND  
TITLE OF INVENTION: METHODS  
FILE REFERENCE: Biv-067.01  
CURRENT APPLICATION NUMBER: US/09/325,256  
CURRENT FILING DATE: 1999-06-03  
PRIOR APPLICATION NUMBER: 60/099,800  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 60/078,935  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/089,685  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/067,423  
PRIOR FILING DATE: 1997-12-03  
PRIOR APPLICATION NUMBER: PCT/US98/25676  
NUMBER OF SEQ ID NOS: 31  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 24  
LENGTH: 396  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-325-256-24

Query Match 13.0%; Score 58.5; DB 4; Length 396;  
Best Local Similarity 30.0%; Pred. No. 21;  
Matches 21; Conservative 6; Mismatches 28; Indels 15; Gaps 2;  
QY 8 CALALLGLALAICSQGAASQPDLDLASRRLL-----ORALAAALPHRSGV--- 52  
| | | | | : | | : | | | | : | | : | | | |



Db 11 CCLALLALPAQSCGPGRGVGRRRYARKQLVPLLYKQFVPGVPERTLGASGPAEGRVARG 70

QY 53 SERWRTFYPN 62

Db 71 SERFRDLVPN 80

RESULT 8

US-09-632-098-2

; Sequence 2, Application US/09632098

; Patent No. 6420154

; GENERAL INFORMATION:

; APPLICANT: Sheppard, Paul O.

; APPLICANT: Baindur, Nand

; APPLICANT: Bishop, Paul D.

; TITLE OF INVENTION: MAMMALIAN ADHESION PROTEASE PEPTIDES

; FILE REFERENCE: 99-39

; CURRENT APPLICATION NUMBER: US/09/632,098

; CURRENT FILING DATE: 2000-08-02

; NUMBER OF SEQ ID NOS: 26

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 2

; LENGTH: 802

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-632-098-2

Query Match

Best Local Similarity 12.9%; Score 58; DB 4; Length 802;

Matches 24; Conservative 5; Mismatches 25; Indels 26; Gaps 4;

QY 10 LALLGLALAICSGAASQPDLDLASRRL-----LQALAAALPHRSGVSE----- 54

Db 617 LDLLGLGLV--EPGTQCGPRMVCQSRRCRKNAFQELQRCLTACHSHGAGLHPSVTSQALV 674

QY 55 -----RWRTFYPNC-PC 65

Db 675 AAWTVALCRLKMTPTSCWPC 694

RESULT 9

US-08-484-105-10

; Sequence 10, Application US/08484105

; Patent No. 5589341

; GENERAL INFORMATION:

; APPLICANT: STILLMAN, Bruce

; APPLICANT: BELL, Stephen P

; APPLICANT: KOBAYASHI, Ryuji

; APPLICANT: RINE, Jasper

; APPLICANT: FOSS, Margit

; APPLICANT: McNALLY, Francis J

; APPLICANT: LAURENSEN, Patricia

; APPLICANT: HERSKOWITZ, Ira

; APPLICANT: LI, Joachim J

; APPLICANT: GAVIN, Kimberly

; TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES

; NUMBER OF SEQUENCES: 24

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT

; STREET: 4 Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94111-4187

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/484,105

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:  
; NAME: Osman Ph.D., Richard Aron  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: A-59032/DJB/RAO  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 494-8700  
; TELEFAX: (415) 494-8771  
; TELEX: 910 277299

; INFORMATION FOR SEQ ID NO: 10:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 479 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-484-105-10

Query Match 12.7%; Score 57; DB 1; Length 479;

Best Local Similarity 30.8%; Pred. No. 41;

Matches 16; Conservative 10; Mismatches 20; Indels 6; Gaps 1;

QY 30 LDLASRRLLQALAAALPHRSGVSEWRWTFYPNCPLRWPRKVKGPQLKAK 81

Db 65 VELVSKPLQLQAIAIT-----VOYKLTLYPNIPTTDYDPLQVEEPFLVK 110

RESULT 10

US-08-484-106-10

; Sequence 10, Application US/08484106

; Patent No. 5614618

; GENERAL INFORMATION:

; APPLICANT: STILLMAN, Bruce

; APPLICANT: BELL, Stephen P

; APPLICANT: KOBAYASHI, Ryuji

; APPLICANT: RINE, Jasper

; APPLICANT: FOSS, Margit

; APPLICANT: McNALLY, Francis J

; APPLICANT: LAURENSEN, Patricia

; APPLICANT: HERSKOWITZ, Ira

; APPLICANT: LI, Joachim J

; APPLICANT: GAVIN, Kimberly

; TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES

; NUMBER OF SEQUENCES: 24

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT

; STREET: 4 Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94111-4187

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/484,106

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Osman Ph.D., Richard Aron

; REGISTRATION NUMBER: 36,627

; REFERENCE/DOCKET NUMBER: A-59032/DJB/RAO

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 494-8700

; TELEFAX: (415) 494-8771

; TELEX: 910 277299

; INFORMATION FOR SEQ ID NO: 10:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 479 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear



MOLECULE TYPE: protein  
US-08-484-106-10

Query Match 12.7%; Score 57; DB 1; Length 479;  
Best Local Similarity 30.8%; Pred. No. 41;  
Matches 16; Conservative 10; Mismatches 20; Indels 6; Gaps 1;

QY 30 LDASRLRLQALAAALPHRSGVSEWRFTYPCNCLRWPRKVKGPQLKAK 81  
DB 65 VELSWKPLQAIART-----VQYKLTLYPNIPTDYDPLQVEEPFLVK 110

RESULT 11

US-09-143-571-29  
; Sequence 29, Application US/09143571  
; Patent No. 6333153

GENERAL INFORMATION:

APPLICANT: FISHEL, Richard  
APPLICANT: GRADIA, Scott  
APPLICANT: ACHARYA, Samir

TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR EFFECTING ADENINE  
TITLE OF INVENTION: NUCLEOTIDE MODULATION OF DNA MISMATCH RECOGNITION  
TITLE OF INVENTION: PROTEINS

FILE REFERENCE: 9855-6U1

CURRENT APPLICATION NUMBER: US/09/143,571

CURRENT FILING DATE: 1998-08-28

EARLIER APPLICATION NUMBER: 60/093,935

EARLIER FILING DATE: 1998-07-23

EARLIER APPLICATION NUMBER: 60/066,977

EARLIER FILING DATE: 1997-11-28

EARLIER APPLICATION NUMBER: 60/057,136

EARLIER FILING DATE: 1997-08-28

NUMBER OF SEQ ID NOS: 36

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 29

LENGTH: 834

TYPE: PRT

ORGANISM: Homo sapiens

US-09-143-571-29

Query Match 12.7%; Score 57; DB 4; Length 834;  
Best Local Similarity 25.7%; Pred. No. 82;  
Matches 28; Conservative 12; Mismatches 39; Indels 30; Gaps 4;

QY 3 VSOIHC-----ALALLGLALAICSQGAASQPDLDLASRRLRLQALAAAL-----46

DB 490 LGLDLCHEIRDQETLLMYQLQCOVLARAVALTRVLDLASRLDVLALASAARDYGYSPRY 549

QY 47 -PHRSGVSEWRFTYPCNCLRWPRKVKGPQLKAK 81

DB 550 SPQVLGVRIQNGRHPMLCARTFVENSTECGDKRGVKVITGPNSSGK 598

RESULT 12

US-08-469-260A-344

; Sequence 344, Application US/08469260A  
; Patent No. 6451578

GENERAL INFORMATION:

APPLICANT: JOHN N. SIMONS

APPLICANT: TAMI J. PILOT-MATIAS

APPLICANT: GEORGE J. DAWSON

APPLICANT: GEORGE G. SCHLAUDER

APPLICANT: SURESH M. DESAI

APPLICANT: THOMAS P. LEARY

APPLICANT: ANTHONY SCOTT MUEHRHOFF

APPLICANT: JAMES C. ERKER

APPLICANT: SHERI L. BUIJK

APPLICANT: ISA K. MUSHAWAR

TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS

TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE

NUMBER OF SEQUENCES: 716

CORRESPONDENCE ADDRESS:

ADDRESSEE: ABBOTT LABORATORIES D377/AP6D

STREET: 100 ABBOTT PARK ROAD  
CITY: ABBOTT PARK  
STATE: IL

COUNTRY: USA

ZIP: 60064-3500

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/469,260A

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/424,550

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: POREBSKI, PRISCILLA E.

REGISTRATION NUMBER: 33,207

REFERENCE/DOCKET NUMBER: 5527 PC.01

TELECOMMUNICATION INFORMATION:

TELEPHONE: 708-937-6365

TELEFAX: 708-938-2623

INFORMATION FOR SEQ ID NO: 344:

SEQUENCE CHARACTERISTICS:

LENGTH: 47 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-469-260A-344

Query Match 12.6%; Score 56.5; DB 4; Length 47;  
Best Local Similarity 41.0%; Pred. No. 2.5;  
Matches 16; Conservative 4; Mismatches 16; Indels 3; Gaps 1;

QY 21 SQGAASQPDLDLASRRLRLQAL--AAALPHRSGVSEWR 56

DB 2 SEGASRPDLRWRTLOLERAVYCACARLQARPGHQRW 40

RESULT 13

US-08-176-427B-4

; Sequence 4, Application US/08176427B

; Patent No. 5789543

GENERAL INFORMATION:

APPLICANT: Ingham, Phillip W.

APPLICANT: McMahon, Andrew P.

APPLICANT: Tabin, Clifford J.

TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing

TITLE OF INVENTION: Proteins and Uses Related Thereto

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD

STREET: 60 State Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII(text)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/176,427B

FILING DATE: 30-DEC-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Vincent, Matthew P.

REGISTRATION NUMBER: 36,709

REFERENCE/DOCKET NUMBER: HMI-006

TELECOMMUNICATION INFORMATION:







Search completed: March 21, 2003, 11:42:33  
Job time : 20.2403 secs

---



GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2003, 11:35:23 ; Search time 53.1502 Seconds  
(without alignments)  
215.607 Million cell updates/sec

Title: US-09-727-739b-17  
Perfect score: 450  
Sequence: 1 MRVSIHICALALLGLALAIC.....RWRPRKVKGPQLKAKEDLER 86

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

| Database :         |                                                  |  |  |
|--------------------|--------------------------------------------------|--|--|
| A_Geneseq_101002.* |                                                  |  |  |
| 1:                 | /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.* |  |  |
| 2:                 | /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.* |  |  |
| 3:                 | /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.* |  |  |
| 4:                 | /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.* |  |  |
| 5:                 | /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.* |  |  |
| 6:                 | /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.* |  |  |
| 7:                 | /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.* |  |  |
| 8:                 | /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.* |  |  |
| 9:                 | /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.* |  |  |
| 10:                | /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.* |  |  |
| 11:                | /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.* |  |  |
| 12:                | /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.* |  |  |
| 13:                | /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.* |  |  |
| 14:                | /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.* |  |  |
| 15:                | /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.* |  |  |
| 16:                | /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.* |  |  |
| 17:                | /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.* |  |  |
| 18:                | /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.* |  |  |
| 19:                | /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.* |  |  |
| 20:                | /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.* |  |  |
| 21:                | /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.* |  |  |
| 22:                | /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.* |  |  |
| 23:                | /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.* |  |  |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID       | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 450   | 100.0       | 111    | AAU07668 | Rainbow trout prep |
| 2          | 353.5 | 78.6        | 115    | AAU07667 | Rainbow trout prep |
| 3          | 85.5  | 19.0        | 114    | AAU07666 | Rainbow trout prep |
| 4          | 79.5  | 17.7        | 116    | AAG03774 | Human secreted pro |
| 5          | 70.5  | 15.7        | 124    | ABG08924 | Novel human diagno |
| 6          | 70    | 15.6        | 517    | AAU02976 | Angiotensin conver |
| 7          | 69.5  | 15.4        | 125    | AAP20029 | Sequence of prepro |
| 8          | 69    | 15.3        | 190    | AAU55408 | Propionibacterium  |
| 9          | 69    | 15.3        | 814    | AAU30069 | Novel human secret |
| 10         | 67.5  | 15.0        | 95     | ABB42770 | Peptide #10276 enc |

|    |      |      |      |    |          |                     |
|----|------|------|------|----|----------|---------------------|
| 11 | 67.5 | 15.0 | 95   | 22 | ABB26063 | Protein #8062 enco  |
| 12 | 67.5 | 15.0 | 95   | 22 | AAM63661 | Human brain expres  |
| 13 | 67.5 | 15.0 | 95   | 22 | AAM76475 | Human bone marrow   |
| 14 | 67.5 | 15.0 | 95   | 22 | AAM20907 | Peptide #7341 enco  |
| 15 | 67.5 | 15.0 | 95   | 22 | AAM36583 | Peptide #10620 enc  |
| 16 | 67.5 | 15.0 | 95   | 23 | ABG45736 | Human peptide enco  |
| 17 | 67.5 | 15.0 | 179  | 22 | AAU62007 | Propionibacterium   |
| 18 | 67   | 14.9 | 502  | 23 | ABP51312 | Human MDDT SEQ ID   |
| 19 | 67   | 14.9 | 1205 | 16 | AAR77363 | Endothelial nitrog  |
| 20 | 67   | 14.9 | 1205 | 17 | AAR93930 | Bovine endothelial  |
| 21 | 66.5 | 14.8 | 283  | 22 | AAU01971 | Human secreted pro  |
| 22 | 65.5 | 14.6 | 146  | 23 | ABG59961 | Human DTHP polype   |
| 23 | 64.5 | 14.3 | 1305 | 20 | AAW79274 | Protein kinase GAK  |
| 24 | 63.5 | 14.1 | 333  | 21 | AAW87092 | Human secreted pro  |
| 25 | 63.5 | 14.1 | 333  | 22 | AAE06069 | Human gene 29 enco  |
| 26 | 63.5 | 14.1 | 333  | 23 | ABG33891 | Human secreted pro  |
| 27 | 63   | 14.0 | 56   | 23 | ABP05384 | Human OREF protein  |
| 28 | 63   | 14.0 | 133  | 22 | ABG27498 | Novel human diagno  |
| 29 | 63   | 14.0 | 283  | 23 | ABB76961 | 4-Hydroxyphenylace  |
| 30 | 63   | 14.0 | 563  | 23 | ABB76956 | 4-Hydroxyphenylace  |
| 31 | 62.5 | 13.9 | 110  | 23 | AAE21871 | Somatostatin-28 (S  |
| 32 | 62.5 | 13.9 | 322  | 22 | AAU44616 | Propionibacterium   |
| 33 | 62   | 13.8 | 119  | 21 | AAV75023 | Neisseria gonorrhoe |
| 34 | 62   | 13.8 | 833  | 20 | AAW94058 | Murine MSH5 (mmsH5  |
| 35 | 61.5 | 13.7 | 68   | 22 | AAU40527 | Propionibacterium   |
| 36 | 61.5 | 13.7 | 92   | 23 | ABP02901 | Human OREF protein  |
| 37 | 61.5 | 13.7 | 333  | 22 | AAU12237 | Human PRO4400 poly  |
| 38 | 61.5 | 13.7 | 333  | 22 | AAW50971 | Human PRO4400 prot  |
| 39 | 61.5 | 13.7 | 375  | 21 | AAV75024 | Neisseria meningit  |
| 40 | 61.5 | 13.7 | 376  | 22 | AAW93305 | Human polypeptide,  |
| 41 | 61.5 | 13.7 | 376  | 23 | AAE14445 | Human drug metabol  |
| 42 | 61.5 | 13.7 | 706  | 23 | ABB92784 | Herbicidally activ  |
| 43 | 61   | 13.6 | 97   | 22 | AAW50353 | Human Rinfintin pr  |
| 44 | 61   | 13.6 | 375  | 21 | AAV75025 | Neisseria meningit  |
| 45 | 61   | 13.6 | 432  | 22 | ABG02228 | Novel human diagno  |

ALIGNMENTS

RESULT 1

AAU07668

ID AAU07668 standard; Protein; 111 AA.

XX AC AAU07668;

DT 04-DEC-2001 (first entry)

XX DE Rainbow trout preprosomatostatin II (PPSS-II'') polypeptide.

KW Rainbow trout; somatostatin; preprosomatostatin; hypersecretion; PPSS-I;  
KW PPSS-II'; endocrine tumour; pituitary gland; glucagonoma; AIDS;  
KW gastroenteropancreatic tissue; acromegaly; gastrinoma; diabetes mellitus;  
KW carcinoid syndrome; cell proliferation; apoptosis; growth hormone;  
KW glucagon; acquired immunodeficiency syndrome; neurological disorder; HIV;  
KW epilepsy; Alzheimer's disease; Huntington's disease; neuroprotective;  
KW neoplasm; metastasis; gene therapy; antidiabetic; nootropic; cytostatic;  
KW anti-human immunodeficiency virus; osteopathic; anticonvulsant.

XX OS Oncochrychus mykiss.

| Key              | Location/Qualifiers             |
|------------------|---------------------------------|
| FT Peptide       | 1..25                           |
| FT Protein       | /note= "Signal peptide"         |
| FT Protein       | /note= "PPSS-II'' pre-sequence" |
| FT Protein       | 26..111                         |
| FT Peptide       | /note= "Mature PPSS-II''"       |
| FT Peptide       | 87..97                          |
| FT Peptide       | /note= "PPSS-II'' pro-sequence" |
| FT Peptide       | 87..111                         |
| FT Cleavage-site | /note= "Prosomatostatin II''"   |
| FT               | 96..97                          |



FT Peptide /note= "Dibasic cleavage site"  
FT 98..111  
FT /note= "SS-14 variant peptide"  
XX  
PN CA2325169-A1.  
XX  
PD 03-JUN-2001.  
XX  
PF 01-DEC-2000; 2000CA-2325169.  
XX  
PR 03-DEC-1999; 99US-0168934.  
XX  
PA (NDSU-) NDSU RES FOUND.  
XX  
PI Sheridan MA, Moore CA, Kittelson JD;  
XX  
DR WPI; 2001-425997/46.  
DR N-PSDB; AAS12935.  
XX  
XX New somatostatin polypeptides derived from Oncorhynchus mykiss, useful  
PT for treating diabetes mellitus, acromegaly, gastrinoma, acquired  
PT immunodeficiency syndrome and neurological disorders -  
XX  
XX Claim 1; Fig 3; 52pp; English.  
CC The invention relates to an Oncorhynchus mykiss somatostatin polypeptide  
CC containing a portion of preprosomatostatin I (PPSS-I) and/or a portion of  
CC preprosomatostatin II (PPSS-II). The protein sequences and their  
CC associated polynucleotides are useful for identifying modified  
CC somatostatin polypeptide which functions as a somatostatin agonist useful  
CC for research, therapeutics or diagnostics, including medical and  
CC veterinary applications. The wild-type somatostatin and its modified  
CC version are useful for treating hypersecretion from endocrine tumours in  
CC the pituitary (e.g. acromegaly) or gastroenteropancreatic tissues (e.g.  
CC gastrinoma, glucagonoma, carcinoid syndrome), to cause tumour shrinkage  
CC through their effects on cell proliferation and apoptosis and as adjuncts  
CC in the treatment of diabetes mellitus via inhibition of growth hormone  
CC and glucagon. In addition, dysfunctional somatostatin secretion is  
CC associated with acquired immunodeficiency syndrome (AIDS) and various  
CC neurological disorders (e.g. epilepsy, Alzheimer's disease and  
CC Huntington's disease) and somatostatin antagonists are effective in the  
CC treatment of such conditions. Nucleic acids encoding the polypeptides are  
CC useful in gene therapy and fusion peptides can be targeted to neoplasms  
CC and their metastases, inhibiting the release of their secretory products.  
CC This sequence represents O. Mykiss PPSS-II', protein.  
CC Note: The features for this sequence are specifically claimed in the  
XX specification.  
SQ Sequence 111 AA;  
Query Match 100.0%; Score 450; DB 22; Length 111;  
Best Local Similarity 100.0%; Pred. No. 3.2e-49;  
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRVSIHCAALLGLALAI CSQGAASQPDLDLSRLLQRLAALPHRSGVSRWRTFY 60  
Db. 1 MRVSIHCAALLGLALAI CSQGAASQPDLDLSRLLQRLAALPHRSGVSRWRTFY 60  
QY 61 PNCPCLRWRPRKVKGPQLKAKEDLER 86  
Db 61 PNCPCLRWRPRKVKGPQLKAKEDLER 86  
RESULT 2  
AAU07667  
ID AAU07667 standard; Protein; 115 AA.  
XX  
AC AAU07667;  
XX  
DT 04-DEC-2001 (first entry)  
XX  
DE Rainbow trout preprosomatostatin II (PPSS-II') polypeptide.  
XX

KW Rainbow trout; somatostatin; preprosomatostatin; hypersecretion; PPSS-I;  
KW PPSS-II'; endocrine tumour; pituitary gland; glucagonoma; AIDS;  
KW gastroenteropancreatic tissue; acromegaly; gastrinoma; diabetes mellitus;  
KW carcinoid syndrome; cell proliferation; apoptosis; growth hormone;  
KW glucagon; acquired immunodeficiency syndrome; neurological disorder; HIV;  
KW epilepsy; Alzheimer's disease; Huntington's disease; neuroprotective;  
KW neoplasm; metastasis; gene therapy; antidiabetic; nootropic; cytostatic;  
KW anti-human immunodeficiency virus; osteopathic; anticonvulsant.  
XX  
OS Oncorhynchus mykiss;  
XX  
XX Location/Qualifiers  
FT Key 1..25  
FT Peptide /note= "Signal peptide"  
FT Protein 1..87  
FT /note= "PPSS-II' pre-sequence"  
FT Protein 26..115  
FT /note= "Mature PPSS-II'"  
FT Misc-difference 74  
FT /note= "Encoded by CAA"  
FT Peptide 88..101  
FT /note= "PPSS-II' pro-sequence"  
FT Peptide 88..115  
FT /note= "Prosomatostatin II'"  
FT Cleavage-site 100..101  
FT /note= "Dibasic cleavage site"  
FT Peptide 102..115  
FT /note= "SS-14 variant peptide"  
XX  
PN CA2325169-A1.  
XX  
PD 03-JUN-2001.  
XX  
PF 01-DEC-2000; 2000CA-2325169.  
XX  
PR 03-DEC-1999; 99US-0168934.  
XX  
PA (NDSU-) NDSU RES FOUND.  
XX  
XX Sheridan MA, Moore CA, Kittelson JD;  
PI WPI; 2001-425997/46.  
DR N-PSDB; AAS12934.  
XX  
XX New somatostatin polypeptides derived from Oncorhynchus mykiss, useful  
PT for treating diabetes mellitus, acromegaly, gastrinoma, acquired  
PT immunodeficiency syndrome and neurological disorders -  
XX  
XX Claim 2; Fig 3; 52pp; English.  
CC The invention relates to an Oncorhynchus mykiss somatostatin polypeptide  
CC containing a portion of preprosomatostatin I (PPSS-I) and/or a portion of  
CC preprosomatostatin II (PPSS-II). The protein sequences and their  
CC associated polynucleotides are useful for identifying modified  
CC somatostatin polypeptide which functions as a somatostatin agonist useful  
CC for research, therapeutics or diagnostics, including medical and  
CC veterinary applications. The wild-type somatostatin and its modified  
CC version are useful for treating hypersecretion from endocrine tumours in  
CC the pituitary (e.g. acromegaly) or gastroenteropancreatic tissues (e.g.  
CC gastrinoma, glucagonoma, carcinoid syndrome), to cause tumour shrinkage  
CC through their effects on cell proliferation and apoptosis and as adjuncts  
CC in the treatment of diabetes mellitus via inhibition of growth hormone  
CC and glucagon. In addition, dysfunctional somatostatin secretion is  
CC associated with acquired immunodeficiency syndrome (AIDS) and various  
CC neurological disorders (e.g. epilepsy, Alzheimer's disease and  
CC Huntington's disease) and somatostatin antagonists are effective in the  
CC treatment of such conditions. Nucleic acids encoding the polypeptides are  
CC useful in gene therapy and fusion peptides can be targeted to neoplasms  
CC and their metastases, inhibiting the release of their secretory products.  
CC This sequence represents O. Mykiss PPSS-II', protein.  
CC Note: The features for this sequence are specifically claimed in the  
XX specification.



SQ Sequence 115 AA;  
Query Match 78.6%; Score 353.5; DB 22; Length 115;  
Best Local Similarity 86.9%; Pred. No. 6e-37;  
Matches 73; Conservative 2; Mismatches 6; Indels 3; Gaps 2;  
QY 1 MRVSIHICALALLGLALATCSQGAASQPDLDLASRRLLQRLAALPHRSGVSRWRTFY 60  
Db 1 MKVCRHICALALLGLALATCSQGAASQPDLDLSRRLQRLAALPHRSGVSRWRTFY 60  
QY 61 PNCPLRWRPRKVKGPQLKAKEDL 84  
Db 61 PNCPL--RPRKVKCP-AGAKEDL 81  
RESULT 3  
AAU07666  
ID AAU07666 standard; Protein; 114 AA.  
XX AC AAU07666;  
XX DT 04-DEC-2001 (first entry)  
XX DE Rainbow trout preprosomatostatin I (PPSS-I) polypeptide.  
KW Rainbow trout; somatostatin; preprosomatostatin; hypersecretion; PPSS-I;  
KW PPSS-II'; endocrine tumour; pituitary gland; glucagonoma; AIDS;  
KW gastroenteropancreatic tissue; acromegaly; gastrinoma; diabetes mellitus;  
KW carcinoid syndrome; cell proliferation; apoptosis; growth hormone;  
KW glucagon; acquired immunodeficiency syndrome; neurological disorder; HIV;  
KW epilepsy; Alzheimer's disease; Huntington's disease; neuroprotective;  
KW neoplasm; metastasis; gene therapy; antidiabetic; nootropic; cytostatic;  
KW anti-human immunodeficiency virus; osteopathic; anticonvulsant.  
XX OS Oncorhynchus mykiss.  
XX FH Key Location/Qualifiers  
FT Peptide 1..24 /note= "Signal peptide"  
FT Protein 1..88 /note= "PPSS-I pre-sequence"  
FT Protein 25..114 /note= "Mature PPSS-I"  
FT Peptide 89..100 /note= "PPSS-I pro-sequence"  
FT Peptide 89..114 /note= "Prosomatostatin I"  
FT Cleavage-site 99..100 /note= "Dibasic cleavage site"  
FT Peptide 101..114 /note= "SS-14 peptide"  
XX CA2325169-A1.  
XX PN 03-JUN-2001.  
XX PD 01-DEC-2000; 2000CA-2325169.  
XX PF 03-DEC-1999; 99US-0168934.  
XX PR (NDSU-) NDSU RES FOUND.  
XX PA Sheridan MA, Moore CA, Kittelson JD;  
XX PI WPI; 2001-425997/46.  
XX DR N-PSDB; AAS12933.  
XX DR New somatostatin polypeptides derived from Oncorhynchus mykiss, useful  
XX PT for treating diabetes mellitus, acromegaly, gastrinoma, acquired  
XX PT immunodeficiency syndrome and neurological disorders.  
XX PS Claim 1; Fig 2; 52pp; English.  
XX

CC The invention relates to an Oncorhynchus mykiss somatostatin polypeptide  
CC containing a portion of preprosomatostatin I (PPSS-I) and/or a portion of  
CC preprosomatostatin II (PPSS-II). The protein sequences and their  
CC associated polynucleotides are useful for identifying modified  
CC somatostatin polypeptide which functions as a somatostatin agonist useful  
CC for research, therapeutics or diagnostics, including medical and  
CC veterinary applications. The wild-type somatostatin and its modified  
CC version are useful for treating hypersecretion from endocrine tumours in  
CC the pituitary (e.g. acromegaly) or gastroenteropancreatic tissues (e.g.  
CC gastrinoma, glucagonoma, carcinoid syndrome), to cause tumour shrinkage  
CC through their effects on cell proliferation and apoptosis and as adjuncts  
CC in the treatment of diabetes mellitus via inhibition of growth hormone  
CC and glucagon. In addition, dysfunctional somatostatin secretion is  
CC associated with acquired immunodeficiency syndrome (AIDS) and various  
CC neurological disorders (e.g. epilepsy, Alzheimer's disease and  
CC Huntington's disease) and somatostatin antagonists are effective in the  
CC treatment of such conditions. Nucleic acids encoding the polypeptides are  
CC useful in gene therapy and fusion peptides can be targeted to neoplasms  
CC and their metastases, inhibiting the release of their secretory products.  
CC This sequence represents O. Mykiss PPSS-I protein.  
CC Note: The features for this sequence are specifically claimed in the  
CC specification.  
XX  
SQ Sequence 114 AA;  
Query Match 19.0%; Score 85.5; DB 22; Length 114;  
Best Local Similarity 52.3%; Pred. No. 0.0065;  
Matches 23; Conservative 5; Mismatches 13; Indels 3; Gaps 1;  
QY 1 MRVSIHICALALLGLALATCSQGAASQPDLDLASRRLLQRLAAL 44  
Db 1 MLSTRVQCALALLSLAIAISSVSAAPS---DAKLRLQLQSLMA 41  
RESULT 4  
AAG03774  
ID AAG03774 standard; Protein; 116 AA.  
XX AC AAG03774;  
XX DT 06-OCT-2000 (first entry)  
XX DE Human secreted protein, SEQ ID NO: 7855;  
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KW gene therapy; chromosome mapping.  
XX OS Homo sapiens.  
XX PN EP1033401-A2.  
XX PD 06-SEP-2000.  
XX PF 21-FEB-2000; 2000EP-0200610.  
XX PR 26-FEB-1999; 99US-0122487.  
XX PA (GEST ) GENSET.  
XX PI Dumas Milne Edwards J, Duclert A, Giordano J;  
XX DR WPI; 2000-500381/45.  
XX DR N-PSDB; AAC03780.  
XX PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
XX PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for  
XX PT diagnostic, forensic, gene therapy and chromosome mapping procedures  
XX PS Claim 13; SEQ ID 7855; 71pp + CD-ROM; English.  
XX CC The present sequence is a polypeptide encoded by one of a large number  
XX CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs  
XX CC were prepared from total human RNAs or polyA+ RNAs derived from 30



different tissues. EST sequences usually correspond mainly to the 3' untranslated region (UTR) of the mRNA because they are often obtained from oligo-dT primed cDNA libraries. Such ESTs are not well suited for isolating cDNA sequences derived from the 5' ends of mRNAs and even in those cases where longer cDNA sequences have been obtained, the full 5' UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and chromosome mapping procedures. They are used to obtain upstream regulatory sequences and to design expression and secretion vectors.

[illegible]

RESULT 5  
ABG08924  
ID ABG08924 standard; Protein; 124 AA.  
XX  
AC  
ABG08924;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE Novel human diagnostic protein #8915.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.

a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic amino acid sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences).

|          |                                                                          |
|----------|--------------------------------------------------------------------------|
| RESULT 6 |                                                                          |
| AAU02976 |                                                                          |
| ID       | AAU02976 standard; Protein; 517 AA.                                      |
| XX       |                                                                          |
| AC       | AAU02976;                                                                |
| XX       |                                                                          |
| DT       | 12-SEP-2001 (first entry)                                                |
| XX       |                                                                          |
| DE       | Angiotensin converting enzyme (ACEV) splice variant protein #76.         |
| XX       |                                                                          |
| KW       | Angiotensin converting enzyme splice variant; ACEV; interleukin 6;       |
| KW       | granulocyte colony stimulating factor receptor; glucagon; hypertrophy;   |
| KW       | platelet-derived endothelial cell growth factor; cardiovascular disease; |
| KW       | cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;       |
| KW       | vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;  |
| KW       | myocardial infarction; coronary arterial thrombosis; renal disease;      |
| KW       | diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;    |
| KW       | multiple sclerosis; immune complex nephritis; deep vein thrombosis;      |
| KW       | nonrheumatoid pulmonary granulomatous disease; endothelial abnormality;  |
| KW       | vascular disorder; asbestosis.                                           |



XX The sequence represents an angiotensin converting enzyme splice variant  
CC (ACEV) polypeptide. The polypeptides of the invention include variants of  
CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,  
CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase  
CC inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal  
CC polypeptide receptor 2. The polypeptides and their associated nucleic  
CC acids are useful for identification of variant sequences and detection of  
CC candidate compounds capable of binding the molecules. The sequences of  
CC the invention can be used in the treatment and diagnosis of various  
CC disorders including cardiovascular diseases such as arteriosclerosis,  
CC myocardial infarction and coronary arterial thrombosis, renal diseases  
CC such as diabetic nephropathy, muscular diseases such as hypertrophy,  
CC immune disorders such as immune complex nephritis, multiple sclerosis,  
CC cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such  
CC as asbestosis and vascular pathologies involving an endothelial  
CC abnormality such as deep vein thrombosis.  
XX

XX SQ Sequence 517 AA;  
Query Match 15.6%; Score 70; DB 22; Length 517;  
Best Local Similarity 37.8%; Pred. No. 3.7;  
Matches 28; Conservative 7; Mismatches 21; Indels 18; Gaps 4;  
QY 14 GLALAICSGAASQPDLDLASRRLLORA-----LAAALPHRSGVSEWRFTYFNCPLR 67  
DB 355 GLARALCSGSPAER-----ROLLPRAREQEBELLAPADGERSGESPSFRLRHP----LP 403  
QY 68 W-RPRKVKGPQLKA 80  
DB 404 FPRPRFPSPRLSA 417

RESULT 7  
AAP20029  
ID AAP20029 standard; Protein; 125 AA.  
XX  
AC AAP20029;  
XX  
XX 16-AUG-2002 (updated)  
DT 14-AUG-1992 (first entry)  
XX  
DE Sequence of preprosomatostatin-2 encoded on pLas2.  
XX  
KW Somatostatin; growth hormone; peptide hormone; secretion.  
X Lophius americanus.  
XX  
XX  
FH Key Location/Qualifiers  
FT Protein 112..125  
FT /label= Somatostatin II  
XX  
PN EP46669-A.  
XX  
PD 03-MAR-1982.  
XX  
XX 21-AUG-1981; 81EP-0303825.  
XX  
XX 25-AUG-1980; 80US-0181046.  
XX  
PA (REGC ) UNIV OF CALIFORNIA.  
XX  
PI Hobart P, Crawford R, Pictet RL, Rutter WJ;  
XX  
XX WPI; 1982-18113E/10.  
DR N-PSDB; AAN20034.  
XX  
XX New somatostatin and precursors - produced by transformed  
PT microorganisms  
XX  
XX Example; Fig 3; 50pp; English.  
PS  
XX The inventors claim preprosomatostatin-1, prosomatostatin-1,

CC preprosomatostatin-2, prosomatostatin-2 and somatostatin-2; and DNA  
CC encoding them. The translation of somatostatin mRNA yields a  
CC precursor (prepro S1) containing a signal peptide which may be  
CC released during the transit into the endoplasmic reticulum, and the  
CC resultant precursor (pro S1) is subsequently cleaved to yield S1  
CC itself. The prepeptide portion of prepro S1 is probably about 20-25  
CC bases long. Translation of pLas2 predicts the sequence of a 125 AA  
CC peptide which surprisingly contains a 14 AA sequence at its carboxy  
CC terminus which differs from S1 by only 2 AAs, and is termed  
CC Somatostatin 2 (S2).  
CC (Updated on 16-AUG-2002 to add missing OS field.)  
XX  
SQ Sequence .125 AA;  
Query Match 15.4%; Score 69.5; DB 3; Length 125;  
Best Local Similarity 36.1%; Pred. No. 0.78;  
Matches 22; Conservative 7; Mismatches 17; Indels 15; Gaps 3;  
QY 6 IHCALALIGLALAICSGAASQ-----PDLDLASRR--LLQRALAALPHRSGVSE 55  
DB 4 IRCPAILALLALVLCGPSVSSQLDREQSDNQDLLELRQHWLLERARSAGL-----LSQE 58  
QY 56 W 56  
DB 59 W 59  
RESULT 8  
AAU55408  
ID AAU55408 standard; Protein; 190 AA.  
XX  
AC AAU55408;  
XX  
DT 27-FEB-2002 (first entry)  
XX  
DE Propionibacterium acnes immunogenic protein #16304.  
XX  
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
KW dermatological; osteopathic; neuroprotectant.  
XX  
OS Propionibacterium acnes.  
XX  
PN WO200181581-A2.  
XX  
PD 01-NOV-2001.  
XX  
XX 20-APR-2001; 2001WO-US12865.  
XX  
XX 21-APR-2000; 2000US-199047P.  
PR 02-JUN-2000; 2000US-208841P.  
PR 07-JUL-2000; 2000US-216747P.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;  
XX  
DR WPI; 2001-616774/71.  
XX N-PSDB; AAS59569.  
XX  
PT Propionibacterium acnes polypeptides and nucleic acids useful for  
PT vaccinating against and diagnosing infections, especially useful for  
PT treating acne vulgaris  
XX  
PS Example 1; SEQ ID No 16603; 1069pp; English.  
XX  
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
CC polypeptides. The proteins and their associated DNA sequences are used in  
CC the treatment, prevention and diagnosis of medical conditions caused by  
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.







CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 95 AA;

Query Match 15.0%; Score 67.5; DB 22; Length 95;  
Best Local Similarity 27.7%; Pred. No. 1;  
Matches 23; Conservative 11; Mismatches 20; Indels 29; Gaps 4;

QY 10 LALLGLALAI-----CSQGAASQPDLDLASRRLLQRALAAALPHRSGVSE 54  
Db 2 LALLYLSLRVNRLLRGNPAYLNTAGSTLGRSSKPSAHFTA-----ASAAASPRSSSLRA 55

QY 55 RW-----RTFYPNCPCLRWRPRK 72  
Db 56 SWSSWDRSWWTRC---AWRPRR 75

RESULT 11  
ABB26063  
ID ABB26063 standard; Protein; 95 AA.  
XX  
AC ABB26063;  
XX

DT 23-JAN-2002 (first entry)  
XX

DE Protein #8062 encoded by probe for measuring heart cell gene expression.  
XX

KW Human; gene expression; heart; microarray; vascular system;  
KW cardiovascular disease; hypertension; cardiac arrhythmia;  
KW congenital heart disease.  
XX

OS Homo sapiens.  
XX

PN WO200157274-A2.  
XX

PD 09-AUG-2001.  
XX

PF 30-JAN-2001; 2001WO-US00666.  
XX

PR 04-FEB-2000; 2000US-0180312.  
XX

PR 26-MAY-2000; 2000US-0207456.  
XX

PR 30-JUN-2000; 2000US-0608408.  
XX

PR 03-AUG-2000; 2000US-0632366.  
XX

PR 21-SEP-2000; 2000US-0234687.  
XX

PR 27-SEP-2000; 2000US-0236359.  
XX

PR 04-OCT-2000; 2000GB-0024263.  
XX

PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX

XX WPI; 2001-488899/53.  
XX

PT Single exon nucleic acid probes for analyzing gene expression in human  
XX hearts.

PS Claim 15; SEQ ID No 27833; 530pp; English.  
XX

CC The present invention relates to single exon nucleic acid probes for  
CC measuring human gene expression in a sample derived from human heart (see  
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such  
CC probe. The probes may be used for predicting, measuring and displaying  
CC gene expression in samples derived from the human heart via microarrays.  
CC By measuring gene expression, the probes are useful for predicting,  
CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
CC human heart and vascular system e.g. cardiovascular disease,  
CC hypertension, cardiac arrhythmias and congenital heart disease.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX

SQ Sequence 95 AA;

Query Match 15.0%; Score 67.5; DB 22; Length 95;  
Best Local Similarity 27.7%; Pred. No. 1;  
Matches 23; Conservative 11; Mismatches 20; Indels 29; Gaps 4;

QY 10 LALLGLALAI-----CSQGAASQPDLDLASRRLLQRALAAALPHRSGVSE 54  
Db 2 LALLYLSLRVNRLLRGNPAYLNTAGSTLGRSSKPSAHFTA-----ASAAASPRSSSLRA 55

QY 55 RW-----RTFYPNCPCLRWRPRK 72  
Db 56 SWSSWDRSWWTRC---AWRPRR 75

RESULT 12  
AAM63661  
ID AAM63661 standard; Protein; 95 AA.  
XX  
AC AAM63661;  
XX

DT 05-NOV-2001 (first entry)  
XX

DE Human brain expressed single exon probe encoded protein SEQ ID NO: 35766.  
XX

KW Human; brain expressed exon; gene expression analysis; probe;  
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
KW epilepsy; cancer.  
XX

OS Homo sapiens.  
XX

PN WO200157275-A2.  
XX

PD 09-AUG-2001.  
XX

PF 30-JAN-2001; 2001WO-US00667.  
XX

PR 04-FEB-2000; 2000US-0180312.  
XX

PR 26-MAY-2000; 2000US-0207456.  
XX

PR 30-JUN-2000; 2000US-0608408.  
XX

PR 03-AUG-2000; 2000US-0632366.  
XX

PR 21-SEP-2000; 2000US-0234687.  
XX

PR 27-SEP-2000; 2000US-0236359.  
XX

PR 04-OCT-2000; 2000GB-0024263.  
XX

PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX

XX WPI; 2001-483446/52.  
XX

PT Single exon nucleic acid probes for analyzing gene expression in human  
XX brains.

PS Example 4; SEQ ID NO: 35766; 650pp + Sequence Listing; English.  
XX

CC The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC brain. They can be used to measure gene expression in brain cell samples,  
CC which may enable the diagnosis and improved treatment of nervous system  
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
CC epilepsy and cancers. The present sequence is a protein encoded by one of  
CC the probes of the invention.  
XX

SQ Sequence 95 AA;

Query Match 15.0%; Score 67.5; DB 22; Length 95;  
Best Local Similarity 27.7%; Pred. No. 1;  
Matches 23; Conservative 11; Mismatches 20; Indels 29; Gaps 4;

QY 10 LALLGLALAI-----CSQGAASQPDLDLASRRLLQRALAAALPHRSGVSE 54  
Db 2 LALLYLSLRVNRLLRGNPAYLNTAGSTLGRSSKPSAHFTA-----ASAAASPRSSSLRA 55



QY 55 RW-----RTFYPNCPCLRWPRK 72  
Db 56 SWSSWDRSWWTRC---AWRPRR 75

RESULT 13  
AAM76475  
ID AAM76475 standard; Protein; 95 AA.  
AC AAM76475;  
XX  
XX  
DT 06-NOV-2001 (first entry)  
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 36781.  
XX  
KW Human; bone marrow expressed exon; gene expression analysis; probe;  
KW microarray; cancer; leukaemia; lymphoma; myeloma.  
XX  
OS Homo sapiens.  
XX  
PN WO200157276-A2.  
XX  
PD 09-AUG-2001.  
30-JAN-2001; 2001WO-US00668.  
04-FEB-2000; 2000US-0180312.  
26-MAY-2000; 2000US-0207456.  
30-JUN-2000; 2000US-0608408.  
03-AUG-2000; 2000US-0632366.  
21-SEP-2000; 2000US-0234687.  
27-SEP-2000; 2000US-0236359.  
04-OCT-2000; 2000GB-0024263.  
(MOLE-) MOLECULAR DYNAMICS INC.  
Penn SG, Hanzel DK, Chen W, Rank DR;  
WPI; 2001-488900/53.  
Human genome-derived single exon nucleic acid probes useful for  
analyzing gene expression in human bone marrow -  
Example 4; SEQ ID NO: 36781; 658pp + Sequence Listing; English.  
The present invention provides a number of single exon nucleic acid  
probes which are derived from genomic sequences expressed in the human  
bone marrow. They can be used to measure gene expression in bone marrow  
samples, which may enable the improved diagnosis and treatment of cancers  
such as lymphoma, leukaemia and myeloma. The present sequence is a  
protein encoded by one of the probes of the invention.

Sequence 95 AA;  
Query Match 15.0%; Score 67.5; DB 22; Length 95;  
Best Local Similarity 27.7%; Pred. No. 1;  
Matches 23; Conservative 11; Mismatches 20; Indels 29; Gaps 4;  
QY 10 LALLGLALAI-----CSQGAASQPDLDLASRRLQLRALAALPHRSGVSE 54  
Db 2 LALLYLSLRVNRLLRGNPAYLNTAGSTLGRSSKPSAHFTA-----ASAAASPRRSLRA 55  
QY 55 RW-----RTFYPNCPCLRWPRK 72  
Db 56 SWSSWDRSWWTRC---AWRPRR 75  
RESULT 14  
AAM20907  
ID AAM20907 standard; Protein; 95 AA.  
AC AAM20907;  
XX

DT 12-OCT-2001 (first entry)  
XX  
DE Peptide #7341 encoded by probe for measuring cervical gene expression.  
XX  
KW Probe; human; microarray; gene expression; cervical epithelial cell;  
KW cervical cancer.  
XX  
OS Homo sapiens.  
XX  
PN WO200157278-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00670.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-488901/53.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human cervical epithelial cells -  
XX  
PS Claim 27; SEQ ID No 25733; 487pp; English.  
XX  
CC The present invention relates to human single exon nucleic acid probes  
CC (SENPs: see AAI10068-AAI28459). The present sequence is a peptide encoded  
CC by one such probe. The SENPs are derived from human Hela cells. The SENPs  
CC can be used to produce a single exon microarray, which can be used for  
CC measuring human gene expression in a sample derived from human cervical  
CC epithelial cells. By measuring gene expression, the probes are therefore  
CC useful in grading and/or staging of diseases of the cervix, notably  
CC cervical cancer.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 95 AA;  
Query Match 15.0%; Score 67.5; DB 22; Length 95;  
Best Local Similarity 27.7%; Pred. No. 1;  
Matches 23; Conservative 11; Mismatches 20; Indels 29; Gaps 4;  
QY 10 LALLGLALAI-----CSQGAASQPDLDLASRRLQLRALAALPHRSGVSE 54  
Db 2 LALLYLSLRVNRLLRGNPAYLNTAGSTLGRSSKPSAHFTA-----ASAAASPRRSLRA 55  
QY 55 RW-----RTFYPNCPCLRWPRK 72  
Db 56 SWSSWDRSWWTRC---AWRPRR 75  
RESULT 15  
AAM36583  
ID AAM36583 standard; Protein; 95 AA.  
XX  
AC AAM36583;  
XX  
DT 17-OCT-2001 (first entry)  
XX  
DE Peptide #10620 encoded by probe for measuring placental gene expression.  
XX  
KW Probe; microarray; human; placenta; antenatal diagnosis;  
KW genetic disorder.



```
XX OS Homo sapiens.
XX PN WO200157272-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00663.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-488897/53.
XX PT Human genome-derived single exon nucleic acid probes useful for
XX PT analyzing gene expression in human placenta.
XX PS Claim 27; SEQ ID No 36852; 654pp; English.
XX CC The present invention relates to single exon nucleic acid probes (SENP:
XX CC see AA131315-AA157546). The present sequence is a peptide encoded by one
XX CC such probe. The probes are useful for producing a microarray for
XX CC predicting, measuring and displaying gene expression in samples derived
XX CC from human placenta. The probes are useful for antenatal diagnosis of
XX CC human genetic disorders.
XX SQ Sequence 95 AA;
Query Match 15.0%; Score 67.5; DB 22; Length 95;
Best Local Similarity 27.7%; Pred No. 1;
Matches 23; Conservative 11; Mismatches 20; Indels 29; Gaps 4;
QY 10 LALLGLALAI-----CSQGAASQPDLDLASRRLQALAAALPHRSGVSE 54
Db 2 LALLYLSLRVNRLLRGNPAYLNTAGSTLGRSSKPSAHFTA-----ASAAASPRRSSLRA 55
Y 55 RW-----RTFYPCPCLRWRPRK 72
Db 56 SWSSWWDNRSWWTRC---AWRPRR 75
```

Search completed: March 21, 2003, 11:38:00  
Job time : 55.1502 secs



GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2003, 11:53:33 ; Search time 2.12446 Seconds  
(without alignments)  
276.816 Million cell updates/sec

Title: US-09-727-739B-18  
Perfect score: 58  
Sequence: 1 SVDNLPPEKK 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 221153 seqs, 53462247 residues

Total number of hits satisfying chosen parameters: 221153

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB pep:\*
- 7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB pep:\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB pep:\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB pep:\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB pep:\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB pep:\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB pep:\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB pep:\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID                     | Description       |
|------------|-------|-------------|--------|------------------------|-------------------|
| 1          | 38    | 65.5        | 33     | 10 US-09-050-516-46    | Sequence 46, Appl |
| 2          | 38    | 65.5        | 197    | 10 US-09-764-860-343   | Sequence 343, App |
| 3          | 35    | 60.3        | 84     | 10 US-09-864-761-43933 | Sequence 43933, A |
| 4          | 35    | 60.3        | 214    | 12 US-10-062-254-156   | Sequence 156, App |
| 5          | 35    | 60.3        | 214    | 12 US-10-062-254-158   | Sequence 158, App |
| 6          | 35    | 60.3        | 398    | 9 US-09-291-417-30     | Sequence 30, Appl |
| 7          | 35    | 60.3        | 409    | 9 US-09-533-029-104    | Sequence 104, App |
| 8          | 35    | 60.3        | 591    | 9 US-09-291-417-103    | Sequence 103, App |
| 9          | 35    | 60.3        | 751    | 10 US-09-815-242-14001 | Sequence 14001, A |
| 10         | 34    | 58.6        | 215    | 10 US-09-050-516-47    | Sequence 47, Appl |
| 11         | 34    | 58.6        | 293    | 10 US-09-815-242-5614  | Sequence 5614, Ap |
| 12         | 34    | 58.6        | 299    | 10 US-09-815-242-10205 | Sequence 10205, A |
| 13         | 34    | 58.6        | 303    | 9 US-09-828-523A-20    | Sequence 20, Appl |
| 14         | 34    | 58.6        | 303    | 10 US-09-815-242-12527 | Sequence 12527, A |
| 15         | 34    | 58.6        | 311    | 9 US-09-828-523A-86    | Sequence 86, Appl |
| 16         | 34    | 58.6        | 528    | 10 US-09-792-200B-22   | Sequence 22, Appl |
| 17         | 34    | 58.6        | 679    | 10 US-09-050-516-42    | Sequence 42, Appl |
| 18         | 33    | 56.9        | 160    | 10 US-09-916-790-35    | Sequence 35, Appl |
| 19         | 33    | 56.9        | 206    | 9 US-09-738-626-6537   | Sequence 6537, Ap |

|    |      |      |      |                        |                   |
|----|------|------|------|------------------------|-------------------|
| 20 | 33   | 56.9 | 265  | 10 US-09-925-122A-1    | Sequence 1, Appli |
| 21 | 33   | 56.9 | 280  | 12 US-10-115-899-5     | Sequence 5, Appli |
| 22 | 33   | 56.9 | 280  | 12 US-10-115-899-8     | Sequence 8, Appli |
| 23 | 33   | 56.9 | 285  | 10 US-09-815-242-4869  | Sequence 4869, Ap |
| 24 | 33   | 56.9 | 334  | 10 US-09-815-242-10798 | Sequence 10798, A |
| 25 | 33   | 56.9 | 345  | 9 US-10-036-729-4      | Sequence 4, Appli |
| 26 | 33   | 56.9 | 547  | 10 US-09-815-242-11982 | Sequence 11982, A |
| 27 | 33   | 56.9 | 683  | 9 US-10-025-380-122    | Sequence 122, App |
| 28 | 33   | 56.9 | 683  | 10 US-09-922-217-122   | Sequence 122, App |
| 29 | 33   | 56.9 | 683  | 10 US-09-833-263-122   | Sequence 122, App |
| 30 | 33   | 56.9 | 764  | 10 US-09-866-582-41    | Sequence 41, Appl |
| 31 | 32   | 55.2 | 77   | 9 US-10-091-483-174    | Sequence 174, App |
| 32 | 32   | 55.2 | 77   | 10 US-09-764-846-174   | Sequence 174, App |
| 33 | 32   | 55.2 | 84   | 9 US-10-091-483-247    | Sequence 247, App |
| 34 | 32   | 55.2 | 84   | 10 US-09-764-846-247   | Sequence 247, App |
| 35 | 32   | 55.2 | 251  | 10 US-09-768-826-35    | Sequence 35, Appl |
| 36 | 32   | 55.2 | 296  | 10 US-09-815-242-13561 | Sequence 13561, A |
| 37 | 32   | 55.2 | 314  | 10 US-09-768-826-54    | Sequence 54, Appl |
| 38 | 32   | 55.2 | 315  | 9 US-10-012-542-297    | Sequence 297, App |
| 39 | 32   | 55.2 | 431  | 9 US-10-012-542-146    | Sequence 146, App |
| 40 | 32   | 55.2 | 678  | 9 US-09-738-626-6025   | Sequence 6025, Ap |
| 41 | 32   | 55.2 | 713  | 10 US-09-802-741A-3    | Sequence 3, Appli |
| 42 | 32   | 55.2 | 715  | 10 US-09-431-226-2     | Sequence 2, Appli |
| 43 | 32   | 55.2 | 1596 | 9 US-09-909-567B-47    | Sequence 47, Appl |
| 44 | 31.5 | 54.3 | 842  | 10 US-09-798-831-8     | Sequence 8, Appli |
| 45 | 31   | 53.4 | 86   | 10 US-09-864-761-46557 | Sequence 46557, A |

ALIGNMENTS

RESULT 1  
US-09-050-516-46

; Sequence 46, Application US/09050516  
; Patent No. US20010010904A1

; GENERAL INFORMATION:

- APPLICANT: BILLING-MEDEL, PATRICIA
- APPLICANT: COHEN, MAURICE
- APPLICANT: COLPITTS, TRACEY L.
- APPLICANT: FRIEDMAN, PAULA N.
- APPLICANT: GORDON, JULIAN
- APPLICANT: GRANADOS, EDWARD N.
- APPLICANT: HAYDEN, MARK
- APPLICANT: HODGES, STEVEN C.
- APPLICANT: KLASS, MICHAEL R.
- APPLICANT: KRATOCHVIL, JON D.
- APPLICANT: ROBERTS-RAPP, LISA
- APPLICANT: RUSSELL, JOHN C.
- APPLICANT: STROUPE, STEPHEN D.

TITLE OF INVENTION: REAGENTS AND METHODS USEFUL FOR DETECTING DISEASES OF THE GASTROINTESTINAL

TITLE OF INVENTION: TRACT

NUMBER OF SEQUENCES: 49

CORRESPONDENCE ADDRESS:

ADDRESSEE: Abbott Laboratories  
STREET: 100 Abbott Park Road  
CITY: Abbott Park  
STATE: IL  
COUNTRY: USA  
ZIP: 60064-3500

COMPUTER-READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/050,516

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/828,855

FILING DATE: 31-MAR-1997

ATTORNEY/AGENT INFORMATION:



NAME: Becker, Cheryl L.  
REGISTRATION NUMBER: 35,441  
REFERENCE/DOCKET NUMBER: 6065 US.P1  
TELEPHONE: 847/935-1729  
TELEFAX: 847/938-2623  
TELEX:  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 33 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: No. US20010010904A1e  
US-09-050-516-46

Query Match 65.5%; Score 38; DB 10; Length 33;  
Best Local Similarity 63.6%; Pred. No. 2.3;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SVDNLPPEPRK 11  
Db 3 SVDSAPPQGRK 13

RESULT 2

US-09-764-860-343  
Sequence 343, Application US/09764860  
Patent No. US20020094953A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PC008  
CURRENT APPLICATION NUMBER: US/09/764,860  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 1198  
SOFTWARE: PatentIn ver. 2.0  
SEQ ID NO 343  
LENGTH: 197  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (53)  
OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (97)  
OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids  
US-09-764-860-343

Query Match 65.5%; Score 38; DB 10; Length 197;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVDNLPPE 7  
Db 174 SVDNLPPE 180

RESULT 3

US-09-864-761-43933  
Sequence 43933, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
FILE REFERENCE: Aemica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761

CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1  
SEQ ID NO 43933  
LENGTH: 84  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC005392.1  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.84  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.56  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.1  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.79  
OTHER INFORMATION: SWISSPROT HIT: Q00888, EVALUE 3.00e-46  
US-09-864-761-43933

Query Match 60.3%; Score 35; DB 10; Length 84;  
Best Local Similarity 54.5%; Pred. No. 23;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 SVDNLPPEPRK 11  
Db 8 TINNLNPRENK 18

RESULT 4

US-10-062-254-156  
Sequence 156, Application US/10062254  
Patent No. US20020138882A1  
GENERAL INFORMATION:  
APPLICANT: Cahoon, Edgar B  
APPLICANT: Cahoon, Rebecca E  
APPLICANT: Falco, Saverio Carl  
APPLICANT: Fang, Yiwen  
APPLICANT: Hantke, Sabine S.  
APPLICANT: Lee, Jian-Ming



APPLICANT: Li, Zhongsen  
APPLICANT: Miao, Guo-Hua  
APPLICANT: Morgante, Michele  
APPLICANT: Niu, Xiping  
APPLICANT: Odell, Joan  
APPLICANT: Rafalski, Antoni  
APPLICANT: Sakai, Hajime  
APPLICANT: Zheng, Peizhong  
APPLICANT: Zhu, Qun

TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved In Plant Metabolism

FILE REFERENCE:  
CURRENT APPLICATION NUMBER: US/10/062,254

CURRENT FILING DATE: 2002-02-01

PRIOR APPLICATION NUMBER: 09/630,346

PRIOR FILING DATE: 2000-07-28

PRIOR APPLICATION NUMBER: 60/146511

PRIOR FILING DATE: 1999-07-30

PRIOR APPLICATION NUMBER: 60/156006

PRIOR FILING DATE: 1999-09-23

PRIOR APPLICATION NUMBER: 60/156899

PRIOR FILING DATE: 1999-09-30

PRIOR APPLICATION NUMBER: 60/157287

PRIOR FILING DATE: 1999-10-01

PRIOR APPLICATION NUMBER: 60/169767

PRIOR FILING DATE: 1999-12-09

PRIOR APPLICATION NUMBER: 60/171054

PRIOR FILING DATE: 1999-12-16

PRIOR APPLICATION NUMBER: 60/172958

PRIOR FILING DATE: 1999-12-21

PRIOR APPLICATION NUMBER: 60/171515

PRIOR FILING DATE: 1999-12-22

PRIOR APPLICATION NUMBER: 60/173535

PRIOR FILING DATE: 1999-12-29

NUMBER OF SEQ ID NOS: 375

SOFTWARE: Microsoft Office 97

SEQ ID NO 156

LENGTH: 214

TYPE: PRT

ORGANISM: Glycine max

FEATURE:

NAME/KEY: UNSURE

LOCATION: (107)

US-10-062-254-156

Query Match 60.3%; Score 35; DB 12; Length 214;

Best Local Similarity 60.0%; Pred. No. 65;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SVDNLPPIR 10

Db 9 SLDHLPSEQ 18

RESULT 5

US-10-062-254-158

Sequence 158, Application US/10062254

Patent No. US2002013882A1

GENERAL INFORMATION:

APPLICANT: Caboon, Edgar B

APPLICANT: Caboon, Rebecca E

APPLICANT: Falco, Saverio Carl

APPLICANT: Fang, Yiwen

APPLICANT: Hantke, Sabine S

APPLICANT: Lee, Jian-Ming

APPLICANT: Li, Zhongsen

APPLICANT: Miao, Guo-Hua

APPLICANT: Morgante, Michele

APPLICANT: Niu, Xiping

APPLICANT: Odell, Joan

APPLICANT: Rafalski, Antoni

APPLICANT: Sakai, Hajime

APPLICANT: Zheng, Peizhong

APPLICANT: Zhu, Qun

TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved In Plant Metabolism

FILE REFERENCE:  
CURRENT APPLICATION NUMBER: US/10/062,254

CURRENT FILING DATE: 2002-02-01

PRIOR APPLICATION NUMBER: 09/630,346

PRIOR FILING DATE: 2000-07-28

PRIOR APPLICATION NUMBER: 60/146511

PRIOR FILING DATE: 1999-07-30

PRIOR APPLICATION NUMBER: 60/156006

PRIOR FILING DATE: 1999-09-23

PRIOR APPLICATION NUMBER: 60/156899

PRIOR FILING DATE: 1999-09-30

PRIOR APPLICATION NUMBER: 60/157287

PRIOR FILING DATE: 1999-10-01

PRIOR APPLICATION NUMBER: 60/169767

PRIOR FILING DATE: 1999-12-09

PRIOR APPLICATION NUMBER: 60/171054

PRIOR FILING DATE: 1999-12-16

PRIOR APPLICATION NUMBER: 60/172958

PRIOR FILING DATE: 1999-12-21

PRIOR APPLICATION NUMBER: 60/171515

PRIOR FILING DATE: 1999-12-22

PRIOR APPLICATION NUMBER: 60/173535

PRIOR FILING DATE: 1999-12-29

NUMBER OF SEQ ID NOS: 375

SOFTWARE: Microsoft Office 97

SEQ ID NO 158

LENGTH: 214

TYPE: PRT

ORGANISM: Glycine max

US-10-062-254-158

Query Match 60.3%; Score 35; DB 12; Length 214;

Best Local Similarity 60.0%; Pred. No. 65;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SVDNLPPIR 10

Db 9 SLDHLPSEQ 18

RESULT 6

US-09-291-417-30

Sequence 30, Application US/09291417A

Publication No. US20030050230A1

GENERAL INFORMATION:

APPLICANT: PLOWMAN, GREGORY

APPLICANT: MARTINEZ, RICARDO

APPLICANT: WHYTE, DAVID

TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES

FILE REFERENCE: 240/300

CURRENT APPLICATION NUMBER: US/09/291,417A

CURRENT FILING DATE: 1999-04-13

EARLIER APPLICATION NUMBER: US 60/081,784

EARLIER FILING DATE: 1998-04-14

NUMBER OF SEQ ID NOS: 147

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 30

LENGTH: 398

TYPE: PRT

ORGANISM: Mammalian (Human) PAKS

US-09-291-417-30

Query Match 60.3%; Score 35; DB 9; Length 398;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 DNLPPR 8

Db 336 DNLPPR 341

RESULT 7



US-09-533-029-104  
; Sequence 104, Application US/09533029  
; Publication No. US20030046723A1  
; GENERAL INFORMATION:  
; APPLICANT: Heard, Jacqueline  
; APPLICANT: Broun, Pierre  
; APPLICANT: Riechmann, Jose-Luis  
; APPLICANT: Keddle, James  
; APPLICANT: Pineda, Omaira  
; APPLICANT: Adam, Luc  
; APPLICANT: Samaha, Raymond  
; APPLICANT: Zhang, James  
; APPLICANT: Yu, Guo-Liang  
; APPLICANT: Ratcliffe, Oliver  
; APPLICANT: Pilgrim, Marsha  
; APPLICANT: Jiang, Cai-Zhong  
; APPLICANT: Reuber, Lynne  
; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES  
; FILE REFERENCE: MBI-010  
; CURRENT APPLICATION NUMBER: US/09/533,029  
; CURRENT FILING DATE: 2000-03-22  
; EARLIER APPLICATION NUMBER: 60/125,814  
; EARLIER FILING DATE: 1999-03-23  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 104  
; LENGTH: 409  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; OTHER INFORMATION: G1034  
US-09-533-029-104

Query Match 60.3%; Score 35; DB 9; Length 409;  
Best Local Similarity 85.7%; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 NLPPRR 10  
|:|:|:|:|  
Db 86 NLPPRR 92

RESULT 8  
US-09-291-417-103  
; Sequence 103, Application US/09291417A  
; Publication No. US20030050230A1  
; GENERAL INFORMATION:  
; APPLICANT: PLOWMAN, GREGORY  
; APPLICANT: MARTINEZ, RICARDO  
; APPLICANT: WHYTE, DAVID  
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES  
; FILE REFERENCE: 240/300  
; CURRENT APPLICATION NUMBER: US/09/291,417A  
; CURRENT FILING DATE: 1999-04-13  
; EARLIER APPLICATION NUMBER: US 60/081,784  
; EARLIER FILING DATE: 1998-04-14  
; NUMBER OF SEQ ID NOS: 147  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 103  
; LENGTH: 591  
; TYPE: PRT  
; ORGANISM: Full Length Mammalian (Human) PAK5hu  
US-09-291-417-103

Query Match 60.3%; Score 35; DB 9; Length 591;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 DNLPPR 8  
|:|:|:|:|  
Db 529 DNLPPR 534

RESULT 9  
US-09-815-242-14001  
; Sequence 14001, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931.  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14001  
; LENGTH: 751  
; TYPE: PRT  
; ORGANISM: Salmonella typhi  
US-09-815-242-14001

Query Match 60.3%; Score 35; DB 10; Length 751;  
Best Local Similarity 85.7%; Pred. No. 2.7e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 DNLPPRE 9  
|:|:|:|:|  
Db 742 DSLPPRE 748

RESULT 10  
US-09-050-516-47  
; Sequence 47, Application US/09050516  
; Patent No. US20010010904A1  
; GENERAL INFORMATION:  
; APPLICANT: BILLING-MEDEL, PATRICIA  
; APPLICANT: COHEN, MAURICE  
; APPLICANT: COLPITTS, TRACEY L.  
; APPLICANT: FRIEDMAN, PAULA N.  
; APPLICANT: GORDON, JULIAN  
; APPLICANT: GRANADOS, EDWARD N.  
; APPLICANT: HAYDEN, MARK  
; APPLICANT: HODGES, STEVEN C.  
; APPLICANT: KLASS, MICHAEL R.  
; APPLICANT: KRATOCHVIL, JON D.  
; APPLICANT: ROBERTS-RAPP, LISA  
; APPLICANT: RUSSELL, JOHN C.  
; APPLICANT: STROUPE, STEPHEN D.  
; TITLE OF INVENTION: REAGENTS AND METHODS USEFUL  
; TITLE OF INVENTION: FOR DETECTING DISEASES OF THE GASTROINTESTINAL  
; NUMBER OF SEQUENCES: 49  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Abbott Laboratories



STREET: 100 Abbott Park Road  
CITY: Abbott Park  
STATE: IL  
COUNTRY: USA  
ZIP: 60064-3500  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/050,516  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA: 08/828,855  
APPLICATION NUMBER: 31-MAR-1997  
FILING DATE: 31-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Becker, Cheryl L.  
REGISTRATION NUMBER: 35,441  
REFERENCE/DOCKET NUMBER: 6065.US.P1  
TELEPHONE: 847/935-1729  
TELEFAX: 847/938-2623  
TELEX:  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 215 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: NO. US20010010904A1e  
US-09-050-516-47

Query Match 58.6%; Score 34; DB 10; Length 215;  
Best Local Similarity 60.0%; Pred. No. 99;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 VDNLPPEPRK 11  
||: ||: ||  
Db 160 VDSAPPGQK 169

RESULT 11  
US-09-815-242-5614  
Sequence 5614, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5614  
LENGTH: 293  
TYPE: PRT  
ORGANISM: Staphylococcus aureus  
US-09-815-242-5614

Query Match 58.6%; Score 34; DB 10; Length 293;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VDNLPP 7  
|||||  
Db 33 VDNLPP 38

RESULT 12  
US-09-815-242-10205  
Sequence 10205, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 10205  
LENGTH: 299  
TYPE: PRT  
ORGANISM: Escherichia coli  
US-09-815-242-10205

Query Match 58.6%; Score 34; DB 10; Length 299;  
Best Local Similarity 77.8%; Pred. No. 1.4e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SVDNLPPE 9  
|||||  
Db 15 STDNLPPE 23

RESULT 13  
US-09-828-523A-20  
Sequence 20, Application US/09828523A  
Patent No. US20020168697A1  
GENERAL INFORMATION:



APPLICANT: The Pharmacia & Upjohn Company  
TITLE OF INVENTION: ANTIMICROBIAL METHODS AND MATERIALS  
FILE REFERENCE: 268.62120101  
CURRENT APPLICATION NUMBER: US/09/828,523A  
CURRENT FILING DATE: 2001-04-06  
PRIOR APPLICATION NUMBER: 60/266,327  
PRIOR FILING DATE: 2000-04-06  
NUMBER OF SEQ ID NOS: 99  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 20  
LENGTH: 303  
TYPE: PRT  
ORGANISM: Staphylococcus aureus  
US-09-828-523A-20

Query Match 58.6%; Score 34; DB 9; Length 303;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VDNLPP 7  
Db 39 VDNLPP 44

SULT 14  
US-09-815-242-12527  
Sequence 12527, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
FILE REFERENCE: ELITRA.011a  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 12527  
LENGTH: 303  
TYPE: PRT  
ORGANISM: Staphylococcus aureus  
US-09-815-242-12527

Query Match 58.6%; Score 34; DB 10; Length 303;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VDNLPP 7  
Db 39 VDNLPP 44

RESULT 15  
US-09-828-523A-86  
Sequence 86, Application US/09828523A  
Patent No. US20020168697A1  
GENERAL INFORMATION:  
APPLICANT: The Pharmacia & Upjohn Company  
TITLE OF INVENTION: ANTIMICROBIAL METHODS AND MATERIALS  
FILE REFERENCE: 268.62120101  
CURRENT APPLICATION NUMBER: US/09/828,523A  
CURRENT FILING DATE: 2001-04-06  
PRIOR APPLICATION NUMBER: 60/266,327  
PRIOR FILING DATE: 2000-04-06  
NUMBER OF SEQ ID NOS: 99  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 86  
LENGTH: 311  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Amino acid sequence encoded by S. aureus coding region cloned  
OTHER INFORMATION: expression in E. coli.  
US-09-828-523A-86

Query Match 58.6%; Score 34; DB 9; Length 311;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VDNLPP 7  
Db 39 VDNLPP 44

Search completed: March 21, 2003, 11:56:31  
Job time: 4.12446 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2003, 11:41:02 ; Search time 2.07725 Seconds  
(without alignments)  
155.808 Million cell updates/sec

Title: US-09-727-739B-18  
Perfect score: 58  
Sequence: 1 SVDNPPRRK 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Issued Patents\_AA:\*

1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*

2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*

3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*

4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*

5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*

6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID                 | Description        |
|------------|-------|-------------|--------|-----------------------|--------------------|
| 1          | 39    | 67.2        | 219    | 4 US-09-247-373B-52   | Sequence 52, Appl  |
| 2          | 39    | 67.2        | 224    | 4 US-09-247-373B-34   | Sequence 34, Appl  |
| 3          | 39    | 67.2        | 248    | 4 US-09-134-001C-5085 | Sequence 5085, Ap  |
| 4          | 37    | 63.8        | 424    | 6 5169835-6           | Patent No. 5169835 |
| 5          | 36    | 62.1        | 273    | 4 US-09-134-001C-3917 | Sequence 3917, Ap  |
| 6          | 36    | 62.1        | 2496   | 4 US-09-125-028-2     | Sequence 2, Appli  |
| 7          | 36    | 62.1        | 2958   | 4 US-08-894-344C-2    | Sequence 2, Appli  |
| 8          | 35    | 60.3        | 230    | 6 5169835-13          | Patent No. 5169835 |
| 9          | 35    | 60.3        | 419    | 6 5169835-2           | Patent No. 5169835 |
| 10         | 35    | 60.3        | 591    | 3 US-09-082-737-2     | Sequence 2, Appli  |
| 11         | 34    | 58.6        | 223    | 3 US-08-857-534-12    | Sequence 12, Appl  |
| 12         | 34    | 58.6        | 223    | 5 PCT-US95-04971-12   | Sequence 12, Appl  |
| 13         | 34    | 58.6        | 229    | 4 US-09-247-373B-48   | Sequence 48, Appl  |
| 14         | 34    | 58.6        | 305    | 4 US-09-134-001C-5632 | Sequence 5632, Ap  |
| 15         | 34    | 58.6        | 361    | 1 US-08-258-261B-3    | Sequence 3, Appli  |
| 16         | 34    | 58.6        | 361    | 1 US-08-456-837-3     | Sequence 3, Appli  |
| 17         | 34    | 58.6        | 361    | 1 US-08-457-342-3     | Sequence 3, Appli  |
| 18         | 34    | 58.6        | 361    | 1 US-08-457-646A-3    | Sequence 3, Appli  |
| 19         | 34    | 58.6        | 361    | 1 US-08-458-076A-3    | Sequence 3, Appli  |
| 20         | 34    | 58.6        | 361    | 1 US-08-457-335A-3    | Sequence 3, Appli  |
| 21         | 34    | 58.6        | 361    | 2 US-08-729-214-3     | Sequence 3, Appli  |
| 22         | 34    | 58.6        | 361    | 3 US-09-028-934-3     | Sequence 3, Appli  |
| 23         | 33    | 56.9        | 180    | 4 US-09-134-001C-5294 | Sequence 5294, Ap  |
| 24         | 33    | 56.9        | 265    | 2 US-08-970-133-1     | Sequence 1, Appli  |
| 25         | 33    | 56.9        | 265    | 4 US-09-294-545-1     | Sequence 1, Appli  |
| 26         | 33    | 56.9        | 280    | 4 US-09-601-478-5     | Sequence 5, Appli  |
| 27         | 33    | 56.9        | 280    | 4 US-09-601-478-8     | Sequence 8, Appli  |

|    |    |      |      |   |                   |                    |
|----|----|------|------|---|-------------------|--------------------|
| 28 | 33 | 56.9 | 345  | 1 | US-08-031-148-4   | Sequence 4, Appli  |
| 29 | 33 | 56.9 | 345  | 3 | US-08-415-838-4   | Sequence 4, Appli  |
| 30 | 33 | 56.9 | 345  | 4 | US-09-205-169-4   | Sequence 4, Appli  |
| 31 | 33 | 56.9 | 415  | 2 | US-08-576-626A-52 | Sequence 52, Appli |
| 32 | 33 | 56.9 | 683  | 1 | US-07-878-960-2   | Sequence 2, Appli  |
| 33 | 33 | 56.9 | 855  | 1 | US-08-344-536-2   | Sequence 2, Appli  |
| 34 | 33 | 56.9 | 855  | 1 | US-08-344-536-5   | Sequence 5, Appli  |
| 35 | 33 | 56.9 | 855  | 3 | US-08-920-562-2   | Sequence 2, Appli  |
| 36 | 33 | 56.9 | 855  | 3 | US-08-920-562-5   | Sequence 5, Appli  |
| 37 | 33 | 56.9 | 2544 | 2 | US-08-576-626A-32 | Sequence 32, Appli |
| 38 | 32 | 55.2 | 275  | 4 | US-09-147-915-4   | Sequence 4, Appli  |
| 39 | 32 | 55.2 | 296  | 3 | US-08-986-769-2   | Sequence 2, Appli  |
| 40 | 32 | 55.2 | 361  | 3 | US-09-028-934-30  | Sequence 30, Appli |
| 41 | 32 | 55.2 | 449  | 1 | US-08-624-663A-2  | Sequence 2, Appli  |
| 42 | 32 | 55.2 | 449  | 2 | US-08-974-565C-1  | Sequence 1, Appli  |
| 43 | 32 | 55.2 | 449  | 3 | US-09-255-748-1   | Sequence 1, Appli  |
| 44 | 32 | 55.2 | 470  | 4 | US-09-004-838-44  | Sequence 44, Appli |
| 45 | 32 | 55.2 | 472  | 4 | US-09-004-838-43  | Sequence 43, Appli |

ALIGNMENTS

RESULT 1  
US-09-247-373B-52  
; Sequence 52, Application US/09247373B  
; Patent No. 6168954  
; GENERAL INFORMATION:  
; APPLICANT: MCGONIGLE, BRIAN  
; APPLICANT: O'KEEFE, DANIEL  
; TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE ENZYMES  
; FILE REFERENCE: CL-1108-A  
; CURRENT APPLICATION NUMBER: US/09/247,373B  
; CURRENT FILING DATE: 1999-02-10  
; PRIOR APPLICATION NUMBER: 08/924,747  
; PRIOR FILING DATE: 1997-09-05  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 52  
; LENGTH: 219  
; TYPE: PRT  
; ORGANISM: SOYBEAN  
US-09-247-373B-52

Query Match 67.2%; Score 39; DB 4; Length 219;  
Best Local Similarity 75.0%; Pred. No. 9.8;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
Qy 3 DNLPPRR 10  
Db 194 DNLPPRDK 201  
|||||

RESULT 2  
US-09-247-373B-34  
; Sequence 34, Application US/09247373B  
; Patent No. 6168954  
; GENERAL INFORMATION:  
; APPLICANT: MCGONIGLE, BRIAN  
; APPLICANT: O'KEEFE, DANIEL  
; TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE ENZYMES  
; FILE REFERENCE: CL-1108-A  
; CURRENT APPLICATION NUMBER: US/09/247,373B  
; CURRENT FILING DATE: 1999-02-10  
; PRIOR APPLICATION NUMBER: 08/924,747  
; PRIOR FILING DATE: 1997-09-05  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 34  
; LENGTH: 224  
; TYPE: PRT  
; ORGANISM: SOYBEAN  
US-09-247-373B-34



```
Query Match      67.2%; Score 39; DB 4; Length 224;
Best Local Similarity 75.0%; Pred. No. 10;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 DNLPPRR 10
DB 202 DNLPPRDK 209

RESULT 3
US-09-134-001C-5085
; Sequence 5085, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5085
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5085

Query Match      67.2%; Score 39; DB 4; Length 248;
Best Local Similarity 70.0%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SVDNLPERR 10
DB 181 SAKNLPKR 190

RESULT 4
5169835-6
; Patent No. 5169835
; APPLICANT: WAI-YEE, CHAN
; TITLE OF INVENTION: PREGNANCY SPECIFIC PROTEINS APPLICATIONS
; NUMBER OF SEQUENCES: 48
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/390,409
; FILING DATE: 07-AUG-1989
; SEQ ID NO: 6
; LENGTH: 424
5169835-6

Query Match      63.8%; Score 37; DB 6; Length 424;
Best Local Similarity 54.5%; Pred. No. 45;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 SVDNLPERR 11
DB 244 TINNLPERR 254

RESULT 5
US-09-134-001C-3917
; Sequence 3917, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
```

```
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3917
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3917

Query Match      62.1%; Score 36; DB 4; Length 273;
Best Local Similarity 60.0%; Pred. No. 42;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 VDNLPERR 11
DB 40 VKHLPKOR 49

RESULT 6
US-09-125-028-2
; Sequence 2, Application US/09125028A
; Patent No. 6190707
; GENERAL INFORMATION:
; APPLICANT: WADOUX, Isabelle
; APPLICANT: COLAVIZZA, Didier
; APPLICANT: LOIEZ, Annie
; TITLE OF INVENTION: NOVEL COLD-SENSITIVE BREAD-MAKING YEASTS
; FILE REFERENCE: levure sensible froid
; CURRENT APPLICATION NUMBER: US/09/125,028A
; CURRENT FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: PCT/FR97/00254
; EARLIER FILING DATE: 1997-02-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2496
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-125-028-2

Query Match      62.1%; Score 36; DB 4; Length 2496;
Best Local Similarity 66.7%; Pred. No. 4.4e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VDNLPERR 10
DB 257 IDVLPKR 265

RESULT 7
US-08-894-344C-2
; Sequence 2, Application US/08894344C
; Patent No. 6172196
; GENERAL INFORMATION:
; APPLICANT: KAWASAKI, Hideki
; APPLICANT: TOKAI, Masaya
; APPLICANT: KIKUCHI, Yasuhiro
; APPLICANT: OUCHI, Kozi
; TITLE OF INVENTION: DNA ENCODING PROTEIN COMPLEMENTING
; TITLE OF INVENTION: YEAST
; TITLE OF INVENTION: LOW TEMPERATURE-SENSITIVE FERMENTABILITY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112-3801
; COMPUTER READABLE FORM:
```



MEDIUM TYPE: Diskette - 3.50 inch, 1440 Kb storage.

COMPUTER: IBM PS/V

OPERATING SYSTEM: MS-DOS Ver3.30

SOFTWARE: PATENT AID Ver1.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/894,344C

FILING DATE: 15-AUGUST-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP343700/95

FILING DATE: 28-DECEMBER-1995

APPLICATION NUMBER: PCT/JP96/03862

FILING DATE: 27-DECEMBER-1996

ATTORNEY/AGENT INFORMATION:

NAME: Perry, Lawrence S.

REGISTRATION NUMBER: 31865

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-218-2100

TELEFAX: 212-218-2200

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 2958 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

ORGANISM: Saccharomyces cerevisiae

STRAIN: X2180-1B

US-08-894-344C-2

Query Match

62.1%; Score 36; DB 4; Length 2958;

Best Local Similarity 66.7%; Pred. No. 5.3e+02;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VDNLPKR 10

|||||

Db 257 IDVLPKR 265

RESULT 8

5169835-13

Patent No. 5169835

APPLICANT: WAI-YEE, CHAN

TITLE OF INVENTION: PREGNANCY SPECIFIC PROTEINS APPLICATIONS

NUMBER OF SEQUENCES: 48

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/390,409

FILING DATE: 07-AUG-1989

SEQ ID NO:13:

LENGTH: 230

5169835-13

Query Match

60.3%; Score 35; DB 6; Length 230;

Best Local Similarity 54.5%; Pred. No. 53;

Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 SVDNLPKR 11

|||||

Db 101 TINLNPENK 111

RESULT 9

5169835-2

Patent No. 5169835

APPLICANT: WAI-YEE, CHAN

TITLE OF INVENTION: PREGNANCY SPECIFIC PROTEINS APPLICATIONS

NUMBER OF SEQUENCES: 48

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/390,409

FILING DATE: 07-AUG-1989

SEQ ID NO:2:

LENGTH: 419

5169835-2

Query Match

60.3%; Score 35; DB 6; Length 419;

Best Local Similarity 54.5%; Pred. No. 1e+02;

Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 SVDNLPKR 11

|||||

Db 245 TINLNPENK 255

RESULT 10

US-09-082-737-2

Sequence 2, Application US/09082737

Patent No. 6013500

GENERAL INFORMATION:

APPLICANT: Minden, Audrey

TITLE OF INVENTION: PAK4; A No. 6013500e1 Gene Encoding A Serine/

TITLE OF INVENTION: Threonine Kinase

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham LLP

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 11230

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/082,737

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 0575/55311

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 278-0400

TELEFAX: (212) 391-0525

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 591 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-082-737-2

Query Match

60.3%; Score 35; DB 3; Length 591;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 DNLPKR 8

|||||

Db 529 DNLPKR 534

RESULT 11

US-08-857-534-12

Sequence 12, Application US/08857534

Patent No. 6087170

GENERAL INFORMATION:

APPLICANT: George W. Kemble

TITLE OF INVENTION: A No. 6087170e1 VZV Gene, Mutant VZV and Immunogenic

TITLE OF INVENTION: Compositions

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooley Godward Castro Huddleson & Tatum

STREET: 5 Palo Alto Square

CITY: Palo Alto

STATE: California

COUNTRY: USA



ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/857,534  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/235,406  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Luann Cseri  
REGISTRATION NUMBER: 31,822  
REFERENCE/DOCKET NUMBER: AVIR-004/000US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-843-5165  
TELEFAX: 415-857-0663  
TELEX: 380816 CooleyPA  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 223 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: Protein  
US-08-857-534-12

Query Match 58.6%; Score 34; DB 3; Length 223;  
Best Local Similarity 100.0%; Pred. No. 78;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 PPRERK 11  
|111111|  
Db 44 PPRERK 49

RESULT 12  
PCT-US95-04971-12  
Sequence 12, Application PC/TUS9504971  
GENERAL INFORMATION:  
APPLICANT: George W. Kemble  
TITLE OF INVENTION: A Novel VZV Gene, Mutant VZV and Immunogenic  
TITLE OF INVENTION: Compositions  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooley Godward Castro Huddleson & Tatum  
STREET: 5 Palo Alto Square  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/04971  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/235,406  
FILING DATE: APRIL 28, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Luann Cseri  
REGISTRATION NUMBER: 31,822  
REFERENCE/DOCKET NUMBER: AVIR-004/000WO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-843-5165  
TELEFAX: 415-857-0663  
TELEX: 380816 CooleyPA

INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 223 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: Protein  
PCT-US95-04971-12  
Query Match 58.6%; Score 34; DB 5; Length 223;  
Best Local Similarity 100.0%; Pred. No. 78;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 6 PPRERK 11  
|111111|  
Db 44 PPRERK 49  
RESULT 13  
US-09-247-373B-48  
Sequence 48, Application US/09247373B  
Patent No. 6168954  
GENERAL INFORMATION:  
APPLICANT: MCGONIGLE, BRIAN  
APPLICANT: O'KEEFE, DANIEL  
TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE ENZYMES  
FILE REFERENCE: CL-1108-A  
CURRENT APPLICATION NUMBER: US/09/247,373B  
CURRENT FILING DATE: 1999-02-10  
PRIOR FILING DATE: 08/924,747  
PRIOR APPLICATION NUMBER: 08/924,747  
NUMBER OF SEQ ID NOS: 56  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 48  
LENGTH: 229  
TYPE: PRT  
ORGANISM: SOYBEAN  
US-09-247-373B-48

Query Match 58.6%; Score 34; DB 4; Length 229;  
Best Local Similarity 62.5%; Pred. No. 80;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 DNLPPRR 10  
|111111|  
Db 200 DSLPPRDK 207

RESULT 14  
US-09-134-001C-5632  
Sequence 5632, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC  
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-08-14  
PRIOR APPLICATION NUMBER: US 60/055,779  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 5632  
LENGTH: 305  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-5632

Query Match 58.6%; Score 34; DB 4; Length 305;  
Best Local Similarity 100.0%; Pred. No. 1.le+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



QY 2 VDNLPP 7  
| | | | |  
Db 42 VDNLPP 47

## RESULT 15

US-08-258-261B-3  
; Sequence 3, Application US/08258261B  
; Patent No. 5639949  
; GENERAL INFORMATION:  
; APPLICANT: Schupp, Thomas  
; APPLICANT: Ligon, James M.  
; APPLICANT: Beck, James Joseph  
; APPLICANT: Hill, Dwight Steven  
; APPLICANT: Ryals, John Andrew  
; APPLICANT: Gaffney, Thomas Deane  
; APPLICANT: Lam, Stephen Ting  
; APPLICANT: Hammer, Phillip E.  
; APPLICANT: Uknes, Scott Joseph  
; TITLE OF INVENTION: Genes for the synthesis of  
; TITLE OF INVENTION: antipathogenic substances  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Ciba-Geigy Corporation  
; STREET: 7 Skyline Drive  
; CITY: Hawthorne  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10532  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/258,261B  
; FILING DATE: 08-JUN-1994  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/457,205  
; FILING DATE: 01-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Elmer, James Scott  
; REGISTRATION NUMBER: 36,129  
; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3  
; TELEPHONE: 919-541-8614  
; TELEFAX: 919-541-8689  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 361 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-258-261B-3

Query Match 58.6%; Score 34; DB 1; Length 361;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 NLPPE 9  
| | | | |  
Db 107 NLPPE 112

Search completed: March 21, 2003, 11:42:35  
Job time : 4.07725 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2003, 11:38:15 ; Search time 12.2318 Seconds  
(without alignments)  
421.131 Million cell updates/sec

Title: US-09-727-739B-19  
Perfect score: 122  
Sequence: 1 MRVSIQHICALALLGLALICSQGA 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues 671580  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_21.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriaph.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB ID | Description         |
|------------|-------|-------|--------|-------|---------------------|
| 1          | 82    | 67.2  | 120    | 13    | Q90Y39 catostomus   |
| 2          | 76    | 62.3  | 114    | 13    | Q90Y42 pantodon bu  |
| 3          | 67    | 54.9  | 114    | 13    | Q90Y41 gnathonemus  |
| 4          | 67    | 54.9  | 115    | 13    | Q90Y43 osteoglossu  |
| 5          | 61    | 50.0  | 116    | 13    | Q90XEL acipenser t  |
| 6          | 56    | 45.9  | 114    | 13    | Q90Y40.chitala chi  |
| 7          | 54    | 44.3  | 216    | 5     | Q96316 aphanartru   |
| 8          | 53    | 43.4  | 434    | 5     | Q96656 penaeus mon  |
| 9          | 52    | 42.6  | 132    | 5     | Q96312 hylurgonotu  |
| 10         | 52    | 42.6  | 133    | 5     | Q963H3 scolychus un |
| 11         | 52    | 42.6  | 146    | 5     | Q96313 xylechinoso  |
| 12         | 52    | 42.6  | 146    | 5     | Q961Z6 dendroctonu  |
| 13         | 52    | 42.6  | 148    | 5     | Q96310 pseudohyles  |
| 14         | 52    | 42.6  | 148    | 5     | Q962A0 hylastes po  |
| 15         | 52    | 42.6  | 149    | 5     | Q96311 dendroctonu  |
| 16         | 52    | 42.6  | 149    | 5     | Q961Z4 dendroctonu  |

|    |    |      |     |    |        |                     |
|----|----|------|-----|----|--------|---------------------|
| 17 | 52 | 42.6 | 149 | 5  | Q961Z0 | Q961Z0 hylurgus mi  |
| 18 | 52 | 42.6 | 150 | 5  | Q961Z2 | Q961Z2 hylurdirecto |
| 19 | 52 | 42.6 | 151 | 5  | Q961Z9 | Q961Z9 dendroctonu  |
| 20 | 52 | 42.6 | 152 | 5  | Q961Z8 | Q961Z8 dendroctonu  |
| 21 | 52 | 42.6 | 153 | 5  | Q961Z3 | Q961Z3 hylurdirecto |
| 22 | 52 | 42.6 | 155 | 5  | Q961Z5 | Q961Z5 dendroctonu  |
| 23 | 52 | 42.6 | 155 | 5  | Q961Z1 | Q961Z1 hylurgonotu  |
| 24 | 52 | 42.6 | 156 | 5  | Q961Z7 | Q961Z7 dendroctonu  |
| 25 | 52 | 42.6 | 158 | 5  | Q963H5 | Q963H5 pityokteine  |
| 26 | 52 | 42.6 | 158 | 5  | Q963H4 | Q963H4 pityophthor  |
| 27 | 52 | 42.6 | 161 | 5  | Q963H7 | Q963H7 hypothernemu |
| 28 | 52 | 42.6 | 164 | 5  | Q963H9 | Q963H9 ctonoxylon   |
| 29 | 52 | 42.6 | 168 | 5  | Q963H8 | Q963H8 haematoloec  |
| 30 | 52 | 42.6 | 168 | 5  | Q963H9 | Q963H9 cerabratalu  |
| 31 | 52 | 42.6 | 168 | 5  | Q963H2 | Q963H2 hymenolepis  |
| 32 | 52 | 42.6 | 169 | 5  | Q963H4 | Q963H4 lacistorhyn  |
| 33 | 52 | 42.6 | 169 | 5  | Q963H5 | Q963H5 stephanosto  |
| 34 | 52 | 42.6 | 170 | 5  | Q963H2 | Q963H2 echinostoma  |
| 35 | 52 | 42.6 | 170 | 5  | Q963H7 | Q963H7 spongilla s  |
| 36 | 52 | 42.6 | 190 | 5  | Q962A1 | Q962A1 hylurgops r  |
| 37 | 52 | 42.6 | 194 | 13 | Q963H8 | Q963H8 gillichthys  |
| 38 | 52 | 42.6 | 216 | 5  | Q963H8 | Q963H8 dendroctonu  |
| 39 | 52 | 42.6 | 219 | 5  | Q963H8 | Q963H8 speleonecte  |
| 40 | 52 | 42.6 | 223 | 5  | Q963H5 | Q963H5 coleobothru  |
| 41 | 52 | 42.6 | 236 | 5  | Q963H4 | Q963H4 stenancylus  |
| 42 | 52 | 42.6 | 238 | 5  | Q963H5 | Q963H5 araucarius   |
| 43 | 52 | 42.6 | 287 | 5  | Q963H2 | Q963H2 taphrorychu  |
| 44 | 52 | 42.6 | 363 | 13 | Q963H6 | Q963H6 amia calva   |
| 45 | 52 | 42.6 | 363 | 13 | Q963H5 | Q963H5 amia calva   |

ALIGNMENTS

RESULT 1  
Q90Y39 ID Q90Y39 PRELIMINARY; PRT; 120 AA.  
AC Q90Y39  
DT 01-DEC-2001 (Tremblrel. 19, Created)  
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
DE Preprosomatostatin.  
OS Catostomus commersoni (White sucker).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Catostomidae; Catostomus.  
OX NCBI\_TaxID=7971;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Al-Mahrouki A.A., Irwin D.M., Youson J.H.;  
RT "Molecular cloning and characterization of white sucker  
preprosomatostatin."  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF292654; AAK97071.1;  
DR InterPro; IPR004250; Somatostatin.  
DR Pfam; PF03002; Somatostatin; 1.  
SQ SEQUENCE 120 AA; 13783 MW; 00828D35263E8805 CRC64;

Query Match 67.2%; Score 82; DB 13; Length 120;  
Best Local Similarity 58.3%; Pred. No. 0.00013;  
Matches 14; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

OY 1 MRVSIQHICALALLGLALICSQGA 24  
||:|||||:|:|:|  
Db 1 MRLCELHCYLLALLGLSLVLCDRGA 24

RESULT 2  
Q90Y42 ID Q90Y42 PRELIMINARY; PRT; 114 AA.  
AC Q90Y42  
DT 01-DEC-2001 (Tremblrel. 19, Created)  
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)



DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Preprosomatostatin.  
 OS Pantodon buchholzi (Butterflyfish).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;  
 OC Osteoglossiformes; Pantodontidae; Pantodon.  
 OX NCBI\_TaxID=8276;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Al-Mahrouti A.A., Irwin D.M., Youson J.H.;  
 RT "Characterization of variant somatostatin cDNAs from several  
 RT osteoglossomorphs: molecular identification and comparative  
 RT analysis.";  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF292651; AAK97068.1;  
 DR InterPro; IPR004250; Somatostatin.  
 DR Pfam; PF03002; Somatostatin; 1.  
 SQ SEQUENCE 114 AA; 12352 MW; 7E3D44CB6A27B12F CRC64;

Query Match 62.3%; Score 76; DB 13; Length 114;  
 Best Local Similarity 52.0%; Pred. No. 0.00095;  
 Matches 13; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Db 1 MRVSIQHICALALLGLALAICSGAA 25  
 1 MKLCQVHCILALLGLVLGMCSSA 25

# RESULT 3 Q90Y41

ID Q90Y41 PRELIMINARY; PRT; 114 AA.  
 AC Q90Y41;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Preprosomatostatin.  
 OS Gnathonemus petersii.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;  
 OC Osteoglossiformes; Mormyridae; Gnathonemus.  
 OX NCBI\_TaxID=42645;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Al-Mahrouti A.A., Irwin D.M., Youson J.H.;  
 RT "Characterization of variant somatostatin cDNAs from several  
 RT osteoglossomorphs: molecular identification and comparative  
 RT analysis.";  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF292652; AAK97069.1;  
 DR InterPro; IPR004250; Somatostatin.  
 DR Pfam; PF03002; Somatostatin; 1.  
 SQ SEQUENCE 114 AA; 12494 MW; 454DA57A309CA8F2 CRC64;

Query Match 54.9%; Score 67; DB 13; Length 114;  
 Best Local Similarity 64.0%; Pred. No. 0.02;  
 Matches 16; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 MRVSIQHICALALLGLALAICSGAA 25  
 1 MLSSRIQCALALLSLALAVSSVSA 25

# RESULT 4 Q90Y43

ID Q90Y43 PRELIMINARY; PRT; 115 AA.  
 AC Q90Y43;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Preprosomatostatin.  
 OS Osteoglossum bicirrhosum (silver arowana).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;

OC Osteoglossiformes; Osteoglossidae; Osteoglossum.  
 OX NCBI\_TaxID=109271;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Al-Mahrouti A.A., Irwin D.M., Youson J.H.;  
 RT "Characterization of variant somatostatin cDNAs from several  
 RT osteoglossomorphs: molecular identification and comparative  
 RT analysis.";  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF292650; AAK97067.1;  
 DR InterPro; IPR004250; Somatostatin.  
 DR Pfam; PF03002; Somatostatin; 1.  
 SQ SEQUENCE 115 AA; 12791 MW; D65FBD7C6F1E4E4D CRC64;

Query Match 54.9%; Score 67; DB 13; Length 115;  
 Best Local Similarity 52.0%; Pred. No. 0.02;  
 Matches 13; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 MRVSIQHICALALLGLALAICSGAA 25  
 1 MKICQIHCTLVLLGLVLGLYCPSAA 25

# RESULT 5 Q90XE1

ID Q90XE1 PRELIMINARY; PRT; 116 AA.  
 AC Q90XE1;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Somatostatin.  
 OS Acipenser transmontanus (White sturgeon).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;  
 OC Acipenser.  
 OX NCBI\_TaxID=7904;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BRAIN;  
 RA Trabucchi M., Tostivint H., Lihmann I., Sollars C., Vallarino M.,  
 RA Dore R.M., Vaudry H.;  
 RT "Polygenic expression of somatostatin in the sturgeon Acipenser  
 RT transmontanus: molecular cloning and distribution of the mRNAs  
 RT encoding two somatostatin precursors.";  
 RL J. Comp. Neurol. 0:0-0(2001).  
 DR EMBL; AF395849; AAL13248.1;  
 DR InterPro; IPR004250; Somatostatin.  
 DR Pfam; PF03002; Somatostatin; 1.  
 SQ SEQUENCE 116 AA; 12616 MW; 72E0C3FF6C80650F CRC64;

Query Match 50.0%; Score 61; DB 13; Length 116;  
 Best Local Similarity 56.0%; Pred. No. 0.15;  
 Matches 14; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 MRVSIQHICALALLGLALAICSGAA 25  
 1 MLSSRLQCALAFLSLALAVSSVSA 25

# RESULT 6 Q90Y40

ID Q90Y40 PRELIMINARY; PRT; 114 AA.  
 AC Q90Y40;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Preprosomatostatin.  
 OS Chitala chitala (clown knife fish).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;  
 OC Osteoglossiformes; Notopteridae; Chitala.  
 OX NCBI\_TaxID=112163;  
 RN [1]



RP SEQUENCE FROM N.A.  
 RA Al-Mahrouki A.A., Irwin D.M., Youson J.H.;  
 RT "Characterization of variant somatostatin cDNAs from several  
 RT osteoglossomorphs: molecular identification and comparative  
 RT analysis."  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF292653; AAK97070.1; -  
 DR InterPro; IPR004250; Somatostatin.  
 DR Pfam; PF03002; Somatostatin; 1.  
 SQ SEQUENCE 114 AA; 12561 MW; 4E3C32F58E34F971 CRC64;

Query Match 45.9%; Score 56; DB 13; Length 114;  
 Best Local Similarity 56.0%; Pred. No. 0.82;  
 Matches 14; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 MRVSQIHCALALGLALAICSGAA 25  
 | : : | | | | | | | : | | |  
 Db 1 MLSTRIQCALALLSLALPVSYYAA 25

## RESULT 7

Q96316 PRELIMINARY; PRT; 216 AA.  
 ID Q96316;  
 AC Q96316;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Enolase (Fragment)  
 GN ENO1NI.  
 OS Aphanarthrum sp. SCRO3.  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;  
 OC Cucujiformia; Phytophaga; Scolytidae; Aphanarthrum.  
 OX NCBI\_TaxID=113463;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RA Farrell B.D., Sequeira A.S., O'Meara B.C., Normark B.B., Chung J.H.,  
 RA Jordal B.H.;  
 RT "The evolution of agriculture in beetles (Curculionidae: Scolytinae  
 RT and Platypodinae).";  
 RL Evolution 0:0-0(2001).  
 DR EMBL; AF375332; AAK54790.1; -  
 DR InterPro; IPR000941; Enolase.  
 DR Pfam; PF00113; enolase; 1.  
 DR ProDom; PD000902; Enolase; 1.  
 FT NON\_TER 1  
 FT NON\_TER 216  
 SQ SEQUENCE 216 AA; 22700 MW; 34E3408775FED864 CRC64;

Query Match 44.3%; Score 54; DB 5; Length 216;  
 Best Local Similarity 58.8%; Pred. No. 2.8;  
 Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALAICSGAA 25  
 | : : | | | | | | | : | | |  
 Db 79 ATAILGVSLAVCKAGAA 95

## RESULT 8

O96656 PRELIMINARY; PRT; 434 AA.  
 ID O96656;  
 AC O96656;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-  
 DE glycerate hydro-lyase).  
 OS Penaeus monodon (Penaeid shrimp).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;  
 OC Penaeidae; Penaeus.  
 OX NCBI\_TaxID=6687;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC TTSUE-ABDOMINAL MUSCLES;  
 RA Boonchuay C., Boonyawan B., Panyim S., Sonthayanon B.;  
 RT "A cDNA sequence of phosphopyruvate hydratase (enolase) from Black  
 RT Tiger Prawns, Penaeus monodon.";  
 RL Asia Pac. J. Mol. Biol. Biotechnol. 7:89-94(1999).  
 CC -1- CATALYTIC ACTIVITY: 2-PHOSPHO-D-GLYCERATE - PHOSPHOENOLPYRUVATE +  
 CC H(2)O.  
 CC -1- COFACTOR: MAGNESIUM IS REQUIRED FOR CATALYSIS AND FOR STABILIZING  
 CC THE DIMER (BY SIMILARITY).  
 CC -1- PATHWAY: GLYCOLYSIS  
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE ENOLASE FAMILY.  
 DR EMBL; AF100985; AAC78141.1; -  
 DR HSSP; P56252; 1PDZ.  
 DR InterPro; IPR000941; Enolase.  
 DR Pfam; PF00113; enolase; 1.  
 DR PRINTS; PR00148; ENOLASE.  
 DR ProDom; PD000902; Enolase; 1.  
 DR TIGRFAMs; TIGR01060; eno; 1.  
 DR PROSITE; PS00164; ENOLASE; 1.  
 DR Glycolysis; Lyase; Magnesium; Pyruvate.  
 KW SEQUENCE 434 AA; 47265 MW; EB575C6FB541ABD3 CRC64;

Query Match 43.4%; Score 53; DB 5; Length 434;  
 Best Local Similarity 64.7%; Pred. No. 7.4;  
 Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALAICSGAA 25  
 | : : | | | | | | | : | | |  
 Db 108 ANAILGVSLAICKAGAA 124

## RESULT 9

Q96312 PRELIMINARY; PRT; 132 AA.  
 ID Q96312;  
 AC Q96312;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Enolase (Fragment)  
 GN ENO1NI.  
 OS Hylurgonotus tuberculatus.  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;  
 OC Cucujiformia; Phytophaga; Scolytidae; Hylurgonotus.  
 OX NCBI\_TaxID=141164;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RA Farrell B.D., Sequeira A.S., O'Meara B.C., Normark B.B., Chung J.H.,  
 RA Jordal B.H.;  
 RT "The evolution of agriculture in beetles (Curculionidae: Scolytinae  
 RT and Platypodinae).";  
 RL Evolution 0:0-0(2001).  
 DR EMBL; AF375337; AAK54795.1; -  
 DR InterPro; IPR000941; Enolase.  
 DR Pfam; PF00113; enolase; 1.  
 DR ProDom; PD000902; Enolase; 1.  
 FT NON\_TER 1  
 FT NON\_TER 132  
 SQ SEQUENCE 132 AA; 13880 MW; BCB814053FFA3471 CRC64;

Query Match 42.6%; Score 52; DB 5; Length 132;  
 Best Local Similarity 58.8%; Pred. No. 3.6;  
 Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALAICSGAA 25  
 | : : | | | | | | | : | | |  
 Db 20 ANAILGVSLAVCKAGAA 36

## RESULT 10



```
Q963H3
ID Q963H3 PRELIMINARY; PRT; 133 AA.
AC Q963H3;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Enolase (Fragment).
GN ENOINI.
OS Scolytus unispinosus.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Cucujiformia; Phytophaga; Scolytidae; Scolytus.
OX NCBI_TaxID=159924;
RN [1]
RP SEQUENCE FROM N.A.
RA Farrell B.D.; Sequeira A.S., O'Meara B.C., Normark B.B., Chung J.H.,
RA Jordal B.H.;
RT "The evolution of agriculture in beetles (Curculionidae: Scolytinae
and Platypodinae).";
RL EMBL; AF375346; AAK54804.1; -.
DR InterPro; IPR000941; Enolase.
DR Pfam; PF00113; enolase; 1.
DR ProDom; PD000902; Enolase; 1.
FT NON_TER 1
FT NON_TER 133 133
SQ SEQUENCE 133 AA; 13879 MW; BF8D60C5D97AC008 CRC64;

Query Match 42.6%; Score 52; DB 5; Length 133;
Best Local Similarity 58.8%; Pred. No. 3.6;
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALAICSGAA 25
| : : : : : | : : :
Db 1 ANAILGVSLAVCKAGAA 17

RESULT 11
Q963I3
ID Q963I3 PRELIMINARY; PRT; 146 AA.
AC Q963I3;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Enolase (Fragment).
GN ENOINI.
OS Xylechinossomus valdivianus.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Cucujiformia; Phytophaga; Scolytidae; Xylechinossomus.
OX NCBI_TaxID=141197;
RN [1]
RP SEQUENCE FROM N.A.
RA Farrell B.D.; Sequeira A.S., O'Meara B.C., Normark B.B., Chung J.H.,
RA Jordal B.H.;
RT "The evolution of agriculture in beetles (Curculionidae: Scolytinae
and Platypodinae).";
RL EMBL; AF375336; AAK54794.1; -.
DR InterPro; IPR000941; Enolase.
DR Pfam; PF00113; enolase; 1.
DR PRINTS; PR01610; CD36ANTIGEN.
DR ProDom; PD000902; Enolase; 1.
FT NON_TER 1
FT NON_TER 146 146
SQ SEQUENCE 146 AA; 15404 MW; E3346CF20ADAF658 CRC64;

Query Match 42.6%; Score 52; DB 5; Length 146;
Best Local Similarity 58.8%; Pred. No. 3.9;
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALAICSGAA 25
| : : : : : | : : :
Db 1 ANAILGVSLAVCKAGAA 17

Q963H3
ID Q963H3 PRELIMINARY; PRT; 133 AA.
AC Q963H3;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Enolase (Fragment).
GN ENOINI.
OS Scolytus unispinosus.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Cucujiformia; Phytophaga; Scolytidae; Scolytus.
OX NCBI_TaxID=159924;
RN [1]
RP SEQUENCE FROM N.A.
RA Farrell B.D.; Sequeira A.S., O'Meara B.C., Normark B.B., Chung J.H.,
RA Jordal B.H.;
RT "The evolution of agriculture in beetles (Curculionidae: Scolytinae
and Platypodinae).";
RL EMBL; AF375346; AAK54804.1; -.
DR InterPro; IPR000941; Enolase.
DR Pfam; PF00113; enolase; 1.
DR ProDom; PD000902; Enolase; 1.
FT NON_TER 1
FT NON_TER 133 133
SQ SEQUENCE 133 AA; 13879 MW; BF8D60C5D97AC008 CRC64;

Query Match 42.6%; Score 52; DB 5; Length 133;
Best Local Similarity 58.8%; Pred. No. 3.6;
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALAICSGAA 25
| : : : : : | : : :
Db 1 ANAILGVSLAVCKAGAA 17

Q963I3
ID Q963I3 PRELIMINARY; PRT; 146 AA.
AC Q963I3;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Enolase (Fragment).
GN ENOINI.
OS Xylechinossomus valdivianus.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Cucujiformia; Phytophaga; Scolytidae; Xylechinossomus.
OX NCBI_TaxID=141197;
RN [1]
RP SEQUENCE FROM N.A.
RA Farrell B.D.; Sequeira A.S., O'Meara B.C., Normark B.B., Chung J.H.,
RA Jordal B.H.;
RT "The evolution of agriculture in beetles (Curculionidae: Scolytinae
and Platypodinae).";
RL EMBL; AF375336; AAK54794.1; -.
DR InterPro; IPR000941; Enolase.
DR Pfam; PF00113; enolase; 1.
DR PRINTS; PR01610; CD36ANTIGEN.
DR ProDom; PD000902; Enolase; 1.
FT NON_TER 1
FT NON_TER 146 146
SQ SEQUENCE 146 AA; 15404 MW; E3346CF20ADAF658 CRC64;

Query Match 42.6%; Score 52; DB 5; Length 146;
Best Local Similarity 58.8%; Pred. No. 3.9;
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALAICSGAA 25
| : : : : : | : : :
Db 1 ANAILGVSLAVCKAGAA 17
```

```
Db 20 ANAILGVSLAVCKAGAA 36

RESULT 12
Q961Z6
ID Q961Z6 PRELIMINARY; PRT; 146 AA.
AC Q961Z6;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Enolase (Fragment).
GN ENOINI.
OS Dendroctonus mexicanus.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Cucujiformia; Phytophaga; Scolytidae; Dendroctonus.
OX NCBI_TaxID=77163;
RN [1]
RP SEQUENCE FROM N.A.
RA Sequeira A.S., Farrell B.D.;
RT "Evolutionary origins of Gondwanan interactions: How old are Araucaria
beetle herbivores?";
RL Biol. J. Linn. Soc. 0:0-0(2001).
DR EMBL; AY040304; AAK77977.1; -.
DR InterPro; IPR000941; Enolase.
DR Pfam; PF00113; enolase; 1.
DR ProDom; PD000902; Enolase; 1.
FT NON_TER 1
FT NON_TER 146 146
SQ SEQUENCE 146 AA; 15194 MW; 28D98F57787CC168 CRC64;

Query Match 42.6%; Score 52; DB 5; Length 146;
Best Local Similarity 58.8%; Pred. No. 3.9;
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALAICSGAA 25
| : : : : : | : : :
Db 13 ANAILGVSLAVCKAGAA 29

RESULT 13
Q963I0
ID Q963I0 PRELIMINARY; PRT; 148 AA.
AC Q963I0;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Enolase (Fragment).
GN ENOINI.
OS Pseudohylesinus nebulosus.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Cucujiformia; Phytophaga; Scolytidae; Pseudohylesinus.
OX NCBI_TaxID=77155;
RN [1]
RP SEQUENCE FROM N.A.
RA Farrell B.D.; Sequeira A.S., O'Meara B.C., Normark B.B., Chung J.H.,
RA Jordal B.H.;
RT "The evolution of agriculture in beetles (Curculionidae: Scolytinae
and Platypodinae).";
RL EMBL; AF375339; AAK54797.1; -.
DR InterPro; IPR000941; Enolase.
DR Pfam; PF00113; enolase; 1.
DR ProDom; PD000902; Enolase; 1.
FT NON_TER 1
FT NON_TER 148 148
SQ SEQUENCE 148 AA; 15554 MW; 3EF31039835D84A0 CRC64;

Query Match 42.6%; Score 52; DB 5; Length 148;
Best Local Similarity 58.8%; Pred. No. 4;
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
```



```
QY 9 ALALLGLALAICSQGAA 25
Db 20 ANAILGVSILAVCKAGAA 36

RESULT 14
Q962A0 ID Q962A0 PRELIMINARY; PRT; 148 AA.
AC Q962A0;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Enolase (Fragment).
GN ENOLNI.
OS Hylastes porculus.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Cucujiformia; Phytophaga; Scolytidae; Hylastes.
OX NCBI_TaxID=141065;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HLY01;
RA "Sequeira A.S., Farrell B.D.;
RT "Evolutionary origins of Gondwanan interactions: How old are Araucaria
RT beetle herbivores?";
RL Biol. J. Linn. Soc. 0:0-0(2001).
DR EMBL; AY040300; AAK77973.1; -.
DR InterPro; IPR000941; Enolase.
DR Pfam; PF00113; enolase; 1.
DR ProDom; PD000902; Enolase; 1.
FT NON_TER 1
FT NON_TER 148 148
FT NON_TER 148 148
SQ SEQUENCE 148 AA; 15404 MW; 633A7D8C7B3D5990 CRC64;
```

Query Match 42.6%; Score 52; DB 5; Length 148;  
Best Local Similarity 58.8%; Pred. No. 4;  
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

```
QY 9 ALALLGLALAICSQGAA 25
Db 13 ANAILGVSILAVCKAGAA 29
```

```
RESULT 15
96311 ID Q96311 PRELIMINARY; PRT; 149 AA.
AC Q96311;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Enolase (Fragment).
GN ENOLNI.
OS Dendroctonus terebrans (black turpentine beetle).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Cucujiformia; Phytophaga; Scolytidae; Dendroctonus.
OX NCBI_TaxID=771172;
RN [1]
RP SEQUENCE FROM N.A.
RA Farrell B.D., Sequeira A.S., O'Meara B.C., Normark B.B., Chung J.H.,
RA Jordal B.H.;
RT "The evolution of agriculture in beetles (Curculionidae: Scolytinae
RT and platypodinae).";
RL Evolution 0:0-0(2001).
DR EMBL; AF375338; AAK54796.1; -.
DR InterPro; IPR000941; Enolase.
DR Pfam; PF00113; enolase; 1.
DR ProDom; PD000902; Enolase; 1.
FT NON_TER 1
FT NON_TER 149 149
FT NON_TER 149 149
SQ SEQUENCE 149 AA; 15570 MW; B09B142A0850D265 CRC64;
```

```
Query Match 42.6%; Score 52; DB 5; Length 149;
Best Local Similarity 58.8%; Pred. No. 4;
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALAICSQGAA 25
Db 19 ANAILGVSILAVCKAGAA 35
```

Search completed: March 21, 2003, 11:40:48  
Job time: 14.2318 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2003, 11:35:23 ; Search time 3.00429 Seconds  
(without alignments)  
345.142 Million cell updates/sec

Title: US-09-727-739B-19  
Perfect score: 122  
Sequence: 1 MRVSIHICALALGLALAICSGAA 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID        | Description         |
|------------|-------|---------------|--------|--------------|---------------------|
| 1          | 110   | 90.2          | 115    | 1 SMS2_ONCMY | Q91194 oncorhynchu- |
| 2          | 74    | 60.7          | 114    | 1 SMSA_CARAU | Q9ygh5 carassius a  |
| 3          | 73    | 59.8          | 120    | 1 SMS2_CARAU | Q9ygh4 carassius a  |
| 4          | 64    | 52.5          | 114    | 1 SMS1_ICTPU | P01171 ictalurus p  |
| 5          | 58    | 47.5          | 115    | 1 SMS1_RANRI | P87384 rana ridibu  |
| 6          | 54    | 44.3          | 433    | 1 ENO_XENLA  | P08734 xenopus lae  |
| 7          | 54    | 44.3          | 815    | 1 AOX1_AERPE | Q9ydx6 aeropyrum p  |
| 8          | 53    | 43.4          | 116    | 1 SMS_CHICK  | P33094 gallus gall  |
| 9          | 53    | 43.4          | 433    | 1 ENO_HOMGA  | P56252 homarus gam  |
| 10         | 52    | 42.6          | 395    | 1 ENO_ALLMI  | P42897 alligator m  |
| 11         | 52    | 42.6          | 433    | 1 ENOA_ALLMI | Q9pvk2 alligator m  |
| 12         | 52    | 42.6          | 433    | 1 ENOA_ANAPL | P19140 anas platyr  |
| 13         | 52    | 42.6          | 433    | 1 ENOA_CHICK | P51913 gallus gall  |
| 14         | 52    | 42.6          | 433    | 1 ENOA_PYTRG | Q9w710 pythion regi |
| 15         | 52    | 42.6          | 433    | 1 ENOA_SCEUN | Q9w712 sceloporu    |
| 16         | 52    | 42.6          | 433    | 1 ENOA_TRASC | Q9w711 trachemys s  |
| 17         | 52    | 42.6          | 433    | 1 ENOB_CHICK | P07322 gallus gall  |
| 18         | 52    | 42.6          | 433    | 1 ENOB_HUMAN | P13929 homo sapien  |
| 19         | 52    | 42.6          | 433    | 1 ENOB_MOUSE | P21550 mus musculu  |
| 20         | 52    | 42.6          | 433    | 1 ENOB_RABIT | P25704 oryctolagus  |
| 21         | 52    | 42.6          | 433    | 1 ENOB_RAT   | P15429 rattus norv  |
| 22         | 52    | 42.6          | 433    | 1 ENOG_HUMAN | P09104 homo sapien  |
| 23         | 52    | 42.6          | 433    | 1 ENOG_MOUSE | P17183 mus musculu  |
| 24         | 52    | 42.6          | 433    | 1 ENOG_RAT   | P07323 rattus norv  |
| 25         | 52    | 42.6          | 434    | 1 ENOG_CHICK | O57391 gallus gall  |
| 26         | 52    | 42.6          | 434    | 1 ENO_LOLPE  | O02654 loligo peal  |
| 27         | 52    | 42.6          | 434    | 1 ENO_SCHJA  | P33676 schistosoma  |
| 28         | 52    | 42.6          | 434    | 1 ENO_SCHMA  | Q27877 schistosoma  |
| 29         | 50    | 41.0          | 439    | 1 ENO1_SCHPO | P40370 schizosacch  |
| 30         | 48    | 39.3          | 431    | 1 ENO_FASHE  | Q27655 fasciola he  |
| 31         | 48    | 39.3          | 433    | 1 ENOA_BOVIN | Q9xsj4 bos taurus   |
| 32         | 48    | 39.3          | 433    | 1 ENOA_HUMAN | P06733 homo sapien  |
| 33         | 48    | 39.3          | 433    | 1 ENOA_MOUSE | P17182 mus musculu  |

|    |      |      |     |              |                    |
|----|------|------|-----|--------------|--------------------|
| 34 | 48   | 39.3 | 433 | 1 ENOA_RAT   | P04764 rattus norv |
| 35 | 47   | 38.5 | 276 | 1 UL34_HSV6U | P52465 human herpe |
| 36 | 47   | 38.5 | 458 | 1 ENOL_HUMAN | Q05524 homo sapien |
| 37 | 46   | 37.7 | 444 | 1 ENO_LYCES  | P26300 lycopersico |
| 38 | 46   | 37.7 | 702 | 1 SPEL_ORYSA | Q9snn0 oryza sativ |
| 39 | 45   | 36.9 | 419 | 1 KDAP_MOUSE | O09043 mus musculu |
| 40 | 45   | 36.9 | 451 | 1 YOE2_STRAT | Q53684 streptomyc  |
| 41 | 44.5 | 36.5 | 116 | 1 SMS_HUMAN  | P04166 homo sapien |
| 42 | 44   | 36.1 | 239 | 1 CLDE_HUMAN | O95500 homo sapien |
| 43 | 44   | 36.1 | 372 | 1 ENO_CHLRE  | P31683 chlamydomon |
| 44 | 44   | 36.1 | 438 | 1 ENO_ASPOR  | Q12560 aspergillus |
| 45 | 44   | 36.1 | 446 | 1 ENO_PLAFA  | Q27727 plasmodium  |

ALIGNMENTS

RESULT 1  
SMS2\_ONCMY STANDARD; PRT; 115 AA.  
AC Q91194;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Somatostatin II precursor [Contains: [Tyr21,Gly24]somatostatin-28;  
DE [Tyr7,Gly10]somatostatin-14].  
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
OX NCBI\_TaxID-8022;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-95354921; PubMed-7628684;  
RA Moore C.A., Kittilson J.D., Dahl S.K., Sheridan M.A.;  
RT "Isolation and characterization of a cDNA encoding for  
RT preprosomatostatin containing [Tyr7, Gly10]-somatostatin-14 from the  
RT endocrine pancreas of rainbow trout, Oncorhynchus mykiss.";  
RL Gen. Comp. Endocrinol. 98:253-261(1995).  
CC -!- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL; U32471; AAC59695.1; -  
CC InterPro; IPR004250; Somatostatin.  
CC Pfam; PF03002; Somatostatin; 1.  
KW Cleavage on pair of basic residues; Hormone; Signal; Multigene family.  
FT SIGNAL 1 18  
FT PROPEP 19 87  
FT PEPTIDE 88 115  
FT PEPTIDE 102 115  
FT DISULFID 104 115  
FT SEQUENCE 115 AA; 12963 MW; 520595025FCA6D91 CRC64;  
DR EMBL; U32471; AAC59695.1; -  
DR InterPro; IPR004250; Somatostatin.  
DR Pfam; PF03002; Somatostatin; 1.  
KW Cleavage on pair of basic residues; Hormone; Signal; Multigene family.  
FT SIGNAL 1 18  
FT PROPEP 19 87  
FT PEPTIDE 88 115  
FT PEPTIDE 102 115  
FT DISULFID 104 115  
FT SEQUENCE 115 AA; 12963 MW; 520595025FCA6D91 CRC64;  
SQ  
Query Match 90.2%; Score 110; DB 1; Length 115;  
Best Local Similarity 88.0%; Pred. No. 4.4e-09;  
Matches 22; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRVSIHICALALGLALAICSGAA 25  
|:|:|||||  
Db 1 MKVCRHICALALGLALAICSGAA 25

RESULT 2  
SMSA\_CARAU



```
ID AC SMSA_CARAU STANDARD; PRT; 114 AA.
DT Q9YGH5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin IA precursor [Contains: Somatostatin-26; Somatostatin-
DE 14].
DE OS Carassius auratus (Goldfish).
DE OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
DE OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
DE OC Cyprinidae; Carassius.
DE OX NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Lin X.-W., Peter R.E.;
RT "Cloning and characterization of cDNAs encoding preprosomatostatin-I
RT and -II from goldfish brain.";
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U40754; AAD09359.1;
DR InterPro; IPR004250; Somatostatin.
DR Pfam; PF03002; Somatostatin; 1.
KW Cleavage on pair of basic residues; Hormone; Signal; Multigene family.
FT SIGNAL 1 24 POTENTIAL.
FT PROPEP 25 88 POTENTIAL.
FT PEPTIDE 89 114 SOMATOSTATIN-26 (POTENTIAL).
FT PEPTIDE 101 114 SOMATOSTATIN-14.
FT DISULFID 103 114 BY SIMILARITY.
SQ SEQUENCE 114 AA; 12574 MW; B5920015E2D272A4 CRC64;

Query Match 60.7%; Score 74; DB 1; Length 114;
Best Local Similarity 64.0%; Pred. No. 0.00062;
Matches 16; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 MRVSOIHCALALLGLALAICSQGA 25
| :| ||||| ||||| ||
Db 1 MLSTRIQCALALLSLALAVCSVSA 25

--SULT 3
32_CARAU
J SMS2_CARAU STANDARD; PRT; 120 AA.
AC Q9YGH4; Q9PTU2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin II precursor [Contains: [Tyr21,Gly24]somatostatin-28;
DE [Tyr7,Gly10]somatostatin-14].
DE OS Carassius auratus (Goldfish).
DE OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
DE OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
DE OC Cyprinidae; Carassius.
DE OX NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Lin X.-W., Peter R.E.;
RT "Cloning and characterization of cDNAs encoding preprosomatostatin-I
RT and -II from goldfish brain.";
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RL [2]
```

```
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Otto C.J., Peter R.E.;
RT "The expression of SRIF mRNA in the brain of goldfish.";
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U60262; AAD09626.1;
DR EMBL; AF025686; AAF15306.1;
DR InterPro; IPR004250; Somatostatin.
DR Pfam; PF03002; Somatostatin; 1.
KW Cleavage on pair of basic residues; Hormone; Signal; Multigene family.
FT SIGNAL 1 23 POTENTIAL.
FT PROPEP 24 92 POTENTIAL.
FT PEPTIDE 93 120 [Tyr21,Gly24]SOMATOSTATIN-28.
FT PEPTIDE 107 120 [Tyr7,Gly10]SOMATOSTATIN-14.
FT DISULFID 109 120 BY SIMILARITY.
FT CONFLICT 51 51 Q -> RW (IN REF. 2).
SQ SEQUENCE 120 AA; 13723 MW; 98957D68011A651A CRC64;

Query Match 59.8%; Score 73; DB 1; Length 120;
Best Local Similarity 54.2%; Pred. No. 0.0009;
Matches 13; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 MRVSOIHCALALLGLALAICSQGA 24
| :| ||||| ||||| ||
Db 1 MRLCELHCYLALLGLSLVLCGRCA 24

RESULT 4
SMS1 ICTPU STANDARD; PRT; 114 AA.
ID SMS1 ICTPU STANDARD; PRT; 114 AA.
AC P01171;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin I precursor [Contains: Somatostatin-14 (SS-14)].
OS Ictalurus punctatus (Channel catfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
OC Ictaluridae; Ictalurus.
OX NCBI_TaxID=7998;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85303576; PubMed=2863931;
RA Dixon-J.E., Andrews P.C.;
RT "Somatostatins of the channel catfish.";
RL Adv. Exp. Med. Biol. 188:19-29(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=82265698; PubMed=6179939;
RA Minth C.D., Taylor W.L., Magazin M.D., Tavlanini M.A., Collier K.J.,
RA Weith H.L., Dixon J.E.;
RT "The structure of cloned DNA complementary to catfish pancreatic
RT somatostatin-14 messenger RNA.";
RL J. Biol. Chem. 257:10372-10377(1982).
RN [3]
RP SEQUENCE OF 82-114 FROM N.A.
RX MEDLINE=82082515; PubMed=6171821;
RA Taylor W.L., Collier K.J., Deschenes R.J., Weith H.L., Dixon J.E.;
RT "Sequence analysis of a cDNA coding for a pancreatic precursor to
RT somatostatin.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:6694-6698(1981).
```



```

RN RP SEQUENCE OF 101-114.
RX MEDLINE=81264223; PubMed=61114953;
RA Andrews P.C., Dixon J.E.;
RT "Isolation and structure of a peptide hormone predicted from a mRNA
sequence. A second somatostatin from the catfish pancreas.";
RL J. Biol. Chem. 256:8267-8270(1981).
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: PANCREAS.
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; M25903; AAA49339.1; -
DR EMBL; V00607; CAA23877.1; -
DR EMBL; V00608; CAA23878.1; -
DR PIR; S00292; RIIDS1.
DR InterPro; IPR004250; Somatostatin.
DR Pfam; PF03002; Somatostatin; 1.
KW Cleavage on pair of basic residues; Hormone; Signal;
KW Multigene family.
FT SIGNAL 1 24 PROBABLE.
FT PEPTIDE 101 114 SOMATOSTATIN-14.
FT DISULFID 103 114
FT CONFLICT 62 62 E -> Q (IN REF. 2).
SQ SEQUENCE 114 AA; 12419 MW; FEE0F2C76F74D99F CRC64;

Query Match 52.5%; Score 64; DB 1; Length 114;
Best Local Similarity 52.0%; Pred. No. 0.017;
Matches 13; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 MRVSIQHCALALLGLALAICSGAA 25
| : : | | | | | : | : | |
Db 1 MPSTRIQCALALLAVALSVCSVSGA 25

RESULT 5
SMS1_RANR1 STANDARD; PRT; 115 AA.
AC P87384; Q9PSI8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin 1 precursor (PSS1) [Contains: Somatostatin-14 (S-I)
(SSS1)].
OS Rana ridibunda (Laughing frog) (Marsh frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranidae; Rana.
OX NCBI_TaxID=8406;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97057290; PubMed=8901629;
RA Tostivint H., Lihmann I., Buchares C., Vieau D., Coulouarn Y.,
Fournier A., Conlon J.M., Vaudry H.;
RT "Occurrence of two somatostatin variants in the frog brain:
characterization of the cDNAs, distribution of the mRNAs, and
receptor-binding affinities of the peptides.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:12605-12610(1996).
RN [2]
RP SEQUENCE OF 102-115.
RC TISSUE=Brain;
RX MEDLINE=93038702; PubMed=1358069;
RA Vaudry H., Chartrel N., Conlon J.M.;
RT "Isolation of [Pro2,Met13]somatostatin-14 and somatostatin-14 from the
frog brain reveals the existence of a somatostatin gene family in a

```

```

tetrapod.";
RL Biochem. Biophys. Res. Commun. 188:477-482(1992).
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; U68136; AAC60093.1; -
DR InterPro; IPR004250; Somatostatin.
DR Pfam; PF03002; Somatostatin; 1.
KW Cleavage on pair of basic residues; Hormone; Multigene family; Signal.
FT SIGNAL 1 24 BY SIMILARITY.
FT PROPEP 25 99 BY SIMILARITY.
FT PEPTIDE 102 115 SOMATOSTATIN-14.
FT DISULFID 104 115 BY SIMILARITY.
SQ SEQUENCE 115 AA; 12691 MW; 349756FEB4ABE213 CRC64;

Query Match 47.5%; Score 58; DB 1; Length 115;
Best Local Similarity 56.0%; Pred. No. 0.12;
Matches 14; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 MRVSIQHCALALLGLALAICSGAA 25
| : : | | | | | | | | | |
Db 1 MQSCRVCALTLLSLALAINISAA 25

RESULT 6
ENO_XENLA STANDARD; PRT; 433 AA.
AC P08734;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-
glycerate hydro-lyase).
GN ENO1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88268812; PubMed=3390159;
RA Segil N., Shrutkowski A., Dworkin M.B., Dworkin-Rastl E.;
RT "Enolase isoenzymes in adult and developing Xenopus laevis and
characterization of a cloned enolase sequence.";
RL Biochem. J. 251:31-39(1988).
CC -1- CATALYTIC ACTIVITY: 2-phospho-D-glycerate -> phosphoenolpyruvate +
H(2)O.
CC -1- COFACTOR: MAGNESIUM IS REQUIRED FOR CATALYSIS AND FOR STABILIZING
THE DIMER.
CC -1- PATHWAY: Glycolysis.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE ENOLASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; Y00718; CAA68706.1; -
DR

```



DR PIR; S00463; NOXL.  
DR HSP; P56252; IPDZ.  
DR InterPro; IPR000941; Enolase.  
DR Pfam; PF00113; enolase; 1.  
DR PRINTS; PR00148; ENOLASE.  
DR ProDom; PD000902; Enolase; 1.  
DR TIGRFAMS; TIGR01060; eno; 1.  
DR PROSITE; PS00164; ENOLASE; 1.  
KW Lyase; Glycolysis; Magnesium.  
FT INIT\_MET 0  
FT ACT\_SITE 157 157 BY SIMILARITY.  
FT METAL 244 244 MAGNESIUM (BY SIMILARITY).  
FT METAL 292 292 MAGNESIUM (BY SIMILARITY).  
FT METAL 317 317 MAGNESIUM (BY SIMILARITY).  
SQ SEQUENCE 433 AA; 47373 MW; 83C922489B530EB4 CRC64;

Query Match 44.3%; Score 54; DB 1; Length 433;  
Best Local Similarity 64.7%; Pred. No. 1.4;  
Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 9 ALALLGLALAICSGAA 25  
Db 107 ANALIGVSLAVCKAGAA 123  
|||||:|:|:|

RESULT 7

AOX1\_AERPE STANDARD; PRT; 815 AA.  
AC Q9YDX6;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Heme-copper oxidase subunit I+III (EC 1.9.3.-).  
GN AOXB OR APE0793.  
OS Aeropyrum pernix.  
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;  
OC Desulfurococaceae; Aeropyrum.  
OX NCBI\_TaxID=56636;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K1;  
RA Wakagi T., Ishikawa R.;  
RT "Heme-copper-oxidase.";  
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K1;  
RX MEDLINE=99310339; PubMed=10382966;  
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,  
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Anka A., Kosugi H.,  
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,  
Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,  
Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,  
Nakamura Y., Nomura N., Sako Y., Kikuchi H.;  
RT "Complete genome sequence of an aerobic-hyper-thermophilic  
Crenarchaeon, Aeropyrum pernix K1.";  
RL DNA Res. 6:83-101(1999).  
CC -1- COFACTOR: TWO HEME GROUPS AND COPPER B (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE HEME-COPPER  
CC RESPIRATORY OXIDASE FAMILY.  
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE CYTOCHROME C  
CC OXIDASE SUBUNIT 3 FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; AB020482; BAA86072.1; --

DR EMBL; AP000060; BAA79771.1; ALT\_INIT.  
DR HSP; P18401; 1FFT.  
DR InterPro; IPR000883; COX1.  
DR InterPro; IPR000298; CytC\_oxdse\_III.  
DR Pfam; PF00115; COX1; 1.  
DR ProDom; PD00510; COX3; 1.  
DR PRINTS; PR01165; CYCOXIDASE1.  
DR ProDom; PD00382; CytC\_oxdse\_III; 1.  
DR PROSITE; PS00077; COX1; 1.  
DR PROSITE; PS0253; COX3; 1.  
KW Oxidoreductase; Heme; Copper; Transmembrane; Respiratory chain;  
KW Complete proteome.  
FT DOMAIN 1 467  
FT DOMAIN 545 815  
FT METAL 70 70 COX1.  
FT METAL 248 248 IRON (HEME A) (PROBABLE).  
FT METAL 252 252 COPPER B (PROBABLE).  
FT METAL 297 297 COPPER B (PROBABLE).  
FT METAL 298 298 COPPER B (PROBABLE).  
FT METAL 383 383 IRON (HEME A3) (PROBABLE).  
FT METAL 385 385 IRON (HEME A) (PROBABLE).  
FT TRANSMEM 26 46  
FT TRANSMEM 71 91 POTENTIAL.  
FT TRANSMEM 105 125 POTENTIAL.  
FT TRANSMEM 157 177 POTENTIAL.  
FT TRANSMEM 197 217 POTENTIAL.  
FT TRANSMEM 242 262 POTENTIAL.  
FT TRANSMEM 281 301 POTENTIAL.  
FT TRANSMEM 314 334 POTENTIAL.  
FT TRANSMEM 339 359 POTENTIAL.  
FT TRANSMEM 380 400 POTENTIAL.  
FT TRANSMEM 419 439 POTENTIAL.  
FT TRANSMEM 463 483 POTENTIAL.  
FT TRANSMEM 580 600 POTENTIAL.  
FT TRANSMEM 637 657 POTENTIAL.  
FT TRANSMEM 683 703 POTENTIAL.  
FT TRANSMEM 708 728 POTENTIAL.  
FT TRANSMEM 736 756 POTENTIAL.  
FT TRANSMEM 758 778 POTENTIAL.  
FT TRANSMEM 791 811 POTENTIAL.  
SQ SEQUENCE 815 AA; 90860 MW; 9E911C57BEF9BBE2 CRC64;

Query Match 44.3%; Score 54; DB 1; Length 815;  
Best Local Similarity 52.2%; Pred. No. 2.4;  
Matches 12; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 2 RVSQIHCAALLGLALAICSGQA 24  
Db 415 RLGKIHFALAMGLVATFLPQFA 437  
|:|:|:|:|:|:|

RESULT 8

ID SMS\_CHICK STANDARD; PRT; 116 AA.  
AC P33094;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Somatostatin precursor [Contains: Somatostatin-28; Somatostatin-14].  
GN SST.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Pancreas;  
RA Nata K., Kobayashi T., Karahashi K., Kato S., Yamamoto H.,  
RA Yonekura H., Okamoto H.;  
RL Submitted (JUN-1991) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.  
CC -1- SUBCELLULAR LOCATION: Secreted.



CC CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.  
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC CC the European Bioinformatics Institute. There are no restrictions on its  
 CC CC use by non-profit institutions as long as its content is in no way  
 CC CC modified and this statement is not removed. Usage by and for commercial  
 CC CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC CC -----

DR EMBL: X60191; CAA42747.1;  
 DR PIR: S20630; S20630.  
 DR InterPro: IPR004250; Somatostatin.  
 DR Pfam: PF03002; Somatostatin; 1.  
 KW Cleavage on pair of basic residues; Hormone; Signal.  
 FT SIGNAL 1 24 BY SIMILARITY.  
 T PROPEP 25 88 BY SIMILARITY.  
 T PEPTIDE 89 116 SOMATOSTATIN-28.  
 FT PEPTIDE 103 116 SOMATOSTATIN-14.  
 FT DISULFID 105 116  
 SQ SEQUENCE 116 AA; 12675 MW; 8A5BB9BDA8A291BA CRC64;

Query Match 43.4%; Score 53; DB 1; Length 116;  
 Best Local Similarity 52.4%; Pred. No. 0.63;  
 Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 5 QIHCALALLGLALAICSQGAA 25  
 Db 5 RLQCALALLSIALAVGVSA 25

RESULT 9  
 ENO\_HOMGA  
 ID ENO\_HOMGA STANDARD; PRT; 433 AA.  
 AC P56252;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-  
 DE glycerate hydro-lyase)  
 OS Homarus gammarus (European lobster) (Homarus vulgaris).  
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;  
 OC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;  
 OC Astacidea; Nephropoidea; Nephropidae; Homarus.  
 NX NCBI\_TaxID=6707;  
 N [1]  
 -1- COFACTOR: MAGNESIUM IS REQUIRED FOR CATALYSIS AND FOR STABILIZING  
 -1- THE DIMER.  
 -1- PATHWAY: Glycolysis.  
 -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
 -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 -1- SIMILARITY: BELONGS TO THE ENOLASE FAMILY.  
 PDB; IPDZ; 14-NOV-95.  
 PDB; IPDZ; 14-NOV-95.  
 InterPro: IPR000941; Enolase.  
 Pfam: PF00113; enolase; 1.  
 PRINTS; PR00148; ENOLASE.  
 ProDom; PD000902; Enolase; 1.  
 TIGRFAMs; TIGR01060; eno; 1.  
 PROSITE; PS00164; ENOLASE; 1.  
 Lyase; Glycolysis; Magnesium; Acetylation; 3D-structure.  
 MOD\_RES 1 157  
 ACT\_SITE 157 157  
 METAL 244 244 MAGNESIUM (BY SIMILARITY).  
 METAL 294 294 MAGNESIUM (BY SIMILARITY).

FT METAL 319 319 MAGNESIUM (BY SIMILARITY).  
 SQ SEQUENCE 433 AA; 47041 MW; CC7DA44B67E9DE3D CRC64;  
 Query Match 43.4%; Score 53; DB 1; Length 433;  
 Best Local Similarity 64.7%; Pred. No. 1.9;  
 Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALAICSQGAA 25  
 Db 107 ANAILGVSLAICKAGAA 123

RESULT 10  
 ENO\_ALLMI  
 ID ENO\_ALLMI STANDARD; PRT; 395 AA.  
 AC P42897;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-  
 DE glycerate hydro-lyase) (Fragment).  
 OS Alligator mississippiensis (American alligator).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Crocodylidae; Alligatorinae; Alligator.  
 NX NCBI\_TaxID=8496;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94195794; PubMed=8146164;  
 RA Hedges S.B.;  
 RT "Molecular evidence for the origin of birds.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:2621-2624(1994).  
 CC -1- CATALYTIC ACTIVITY: 2-phospho-D-glycerate - phosphoenolpyruvate +  
 CC H(2)O.  
 CC -1- COFACTOR: MAGNESIUM IS REQUIRED FOR CATALYSIS AND FOR STABILIZING  
 CC THE DIMER (BY SIMILARITY).  
 CC -1- PATHWAY: Glycolysis.  
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO THE ENOLASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: L28078; AAA53671.1;  
 DR HSSP; P56252; IPDZ.  
 DR InterPro: IPR000941; Enolase.  
 DR Pfam; PF00113; enolase; 1.  
 DR ProDom; PD000902; Enolase; 1.  
 DR TIGRFAMs; TIGR01060; eno; 1.  
 DR PROSITE; PS00164; ENOLASE; 1.  
 KW Lyase; Glycolysis; Magnesium.  
 FT NON\_TER 1 1  
 FT ACT\_SITE 136 136 BY SIMILARITY.  
 FT METAL 223 223 MAGNESIUM (BY SIMILARITY).  
 FT METAL 271 271 MAGNESIUM (BY SIMILARITY).  
 FT METAL 296 296 MAGNESIUM (BY SIMILARITY).  
 FT NON\_TER 395 395  
 SQ SEQUENCE 395 AA; 42884 MW; B43E91228E9110B5 CRC64;

Query Match 42.6%; Score 52; DB 1; Length 395;  
 Best Local Similarity 58.8%; Pred. No. 2.5;  
 Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALAICSQGAA 25  
 Db 86 ANAILGVSLAICKAGAA 102



```
RESULT 11
ENOA_ALLMI
ID ENOA_ALLMI STANDARD; PRT; 433 AA.
AC Q9PVK2;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Alpha enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase)
DE (Phosphopyruvate hydratase).
OS Alligator mississippiensis (American alligator).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Crocodylidae; Alligatorinae; Alligator.
OX NCBI_TaxID=8496;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RX MEDLINE=99439677; PubMed=10508547;
RA Mannen H., Li S.S.-L.;
RT "Molecular evidence for a clade of turtles.";
RL Mol. Phylogenet. Evol. 13:144-148(1999).
CC -!- CATALYTIC ACTIVITY: 2-phospho-D-glycerate - phosphoenolpyruvate +
CC H(2)O.
CC -!- COFACTOR: MAGNESIUM IS REQUIRED FOR CATALYSIS AND FOR STABILIZING
CC THE DIMER (BY SIMILARITY).
CC -!- PATHWAY: Glycolysis.
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE ENOLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF072586; AAD41643.1;
CC HSSP; P56252; LPDZ.
CC InterPro; IPR000941; Enolase.
CC Pfam; PF00113; enolase; 2.
CC PRINTS; PR00148; ENOLASE.
CC ProDom; PD000902; Enolase; 1.
CC TIGRFAMS; TIGR01060; eno; 1.
CC PROSITE; PS00164; ENOLASE; 1.
CC Lyase; Glycolysis; Magnesium; Multigene family.
KW INIT_MET 0
FT METAL 244 244 MAGNESIUM (BY SIMILARITY).
FT METAL 292 292 MAGNESIUM (BY SIMILARITY).
FT METAL 317 317 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 433 AA; 47190 MW; CD3DD59FB1EBB3DA CRC64;

Query Match 42.6%; Score 52; DB 1; Length 433;
Best Local Similarity 58.8%; Pred. No. 2.7;
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALAICSGGAA 25
| : : : : : : : : : :
Db 107 ANAILGVSLAVCKAGAA 123

RESULT 12
ENOA_ANAPL
ID ENOA_ANAPL STANDARD; PRT; 433 AA.
AC P19140;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Alpha enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (Tau-
DE crystallin).
OS Anas platyrhynchos (Domestic duck).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anas.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White leghorn; Tissue=Kidney;
RX MEDLINE=95355305; PubMed=7629021;
```

```
OX NCBI_TaxID=8839;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryonic lens;
RX MEDLINE=89079778; PubMed=2462567;
RA Wistow G.J., Lietman T., Williams L.A., Stapel S.O., de Jong W.W.,
RA Horwitz J., Platiogorsky J.;
RT "Tau-crystallin/alpha-enolase: one gene encodes both an enzyme and a
RT lens structural protein.";
RL J. Cell Biol. 107:2729-2736(1988).
CC -!- FUNCTION: BOTH AN ENZYME AND A LENS STRUCTURAL PROTEIN.
CC -!- CATALYTIC ACTIVITY: 2-phospho-D-glycerate - phosphoenolpyruvate +
CC H(2)O.
CC -!- COFACTOR: MAGNESIUM IS REQUIRED FOR CATALYSIS AND FOR STABILIZING
CC THE DIMER.
CC -!- PATHWAY: Glycolysis.
CC -!- SUBUNIT: HOMODIMER.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE ENOLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M20749; AAA49218.1;
CC EMBL; X14195; CAA32409.1;
CC PIR; A32132; A32132.
CC HSSP; P56252; LPDZ.
CC InterPro; IPR000941; Enolase.
CC Pfam; PF00113; enolase; 1.
CC PRINTS; PR00148; ENOLASE.
CC ProDom; PD000902; Enolase; 1.
CC TIGRFAMS; TIGR01060; eno; 1.
CC PROSITE; PS00164; ENOLASE; 1.
CC Lyase; Glycolysis; Magnesium; Eye lens protein.
KW INIT_MET 0
FT ACT_SITE 157 157 BY SIMILARITY.
FT METAL 244 244 MAGNESIUM (BY SIMILARITY).
FT METAL 292 292 MAGNESIUM (BY SIMILARITY).
FT METAL 317 317 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 433 AA; 47108 MW; 1AED78B08A66E84D CRC64;

Query Match 42.6%; Score 52; DB 1; Length 433;
Best Local Similarity 58.8%; Pred. No. 2.7;
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALAICSGGAA 25
| : : : : : : : : : :
Db 107 ANAILGVSLAVCKAGAA 123

RESULT 13
ENOA_CHICK
ID ENOA_CHICK STANDARD; PRT; 433 AA.
AC P51913;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Alpha enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase)
DE (Phosphopyruvate hydratase).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White leghorn; Tissue=Kidney;
RX MEDLINE=95355305; PubMed=7629021;
```







```
CC -----
DR EMBL; AF072587; AAD41644.1;
DR HSP; P56252; IPDZ.
DR InterPro; IPR000941; Enolase.
DR Pfam; PF00113; enolase; 2.
DR PRINTS; PR00148; ENOLASE.
DR ProDom; PD000902; Enolase; 1.
DR TIGRFAMs; TIGR01060; eno; 1.
DR PROSITE; PS00164; ENOLASE; 1.
KW Lyase; Glycolysis; Magnesium; Multigene family.
FT INIT_MET 0 0 BY SIMILARITY.
FT METAL 244 244 MAGNESIUM (BY SIMILARITY).
FT METAL 292 292 MAGNESIUM (BY SIMILARITY).
FT METAL 317 317 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 433 AA; 47362 MW; 7CD89EB405529301 CRC64;

Query Match 42.6%; Score 52; DB 1; Length 433;
Best Local Similarity 58.8%; Pred. No. 2.7;
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 9 ALALLGLALAICSQGAA 25
Db 107 ANAILGVSLAVCKAGAA 123
```

Search completed: March 21, 2003, 11:38:45  
Job time : 5.00429 secs











Db 418 RLGIHFALAMGLVALTELPQFA 440

RESULT 7

S20630  
somatostatin precursor - chicken  
C;Species: Gallus gallus (chicken)  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C;Accession: S20630  
R;Nata, K.; Kobayashi, T.; Karahashi, K.; Kato, S.; Yamamoto, H.; Yonekura, H.; Okamoto, S.  
submitted to the EMBL Data Library, June 1991  
A;Description: Nucleotide sequence determination of chicken somatostatin precursor cDNA.  
A;Reference number: S20630  
A;Accession: S20630  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-116 <NAT>  
A;Cross-references: EMBL:X60191; NID:g62985; PIDN:CAA42747.1; PID:g62986  
C;Superfamily: somatostatin

Query Match 43.4%; Score 53; DB 1; Length 116;  
Best Local Similarity 52.4%; Pred. No. 1.4;  
Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 5 QIHCALALLGLALAICSQGAA 25

Db 5 RLQCALALLSIALAVGTVSAA 25

RESULT 8

I50026  
phosphopyruvate hydratase (EC 4.2.1.11) alpha - American alligator (fragment)  
N;Alternate names: alpha-enolase  
C;Species: Alligator mississippiensis (American alligator)  
C;Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 22-Jun-1999  
C;Accession: I50026  
R;Hedges, S.B.  
Proc. Natl. Acad. Sci. U.S.A. 91, 2621-2624, 1994  
A;Title: Molecular evidence for the origin of birds.  
A;Reference number: A53470; MUID:94195794; PMID:8146164  
A;Accession: I50026  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown; translated  
A;Molecule type: mRNA  
A;Residues: 1-395 <HED>  
A;Cross-references: GB:L28078; NID:g472796; PIDN:AAA53671.1; PID:g472797  
C;Superfamily: enolase  
C;Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 42.6%; Score 52; DB 2; Length 395;  
Best Local Similarity 58.8%; Pred. No. 5.3;  
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALAICSQGAA 25

Db 86 ANAILGVSLAVCKAGAA 102

RESULT 9

A53665  
phosphopyruvate hydratase (EC 4.2.1.11) - liver fluke  
C;Species: Fasciola hepatica (liver fluke)  
C;Date: 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 22-Jun-1999  
C;Accession: A53665  
R;Davis, R.E.; Singh, H.; Botka, C.; Hardwick, C.; Ashraf el Meanawy, M.; Villanueva, J.  
J. Biol. Chem. 269, 20026-20030, 1994  
A;Title: RNA trans-splicing in Fasciola hepatica. Identification of a spliced leader (SL)  
A;Reference number: A53665; MUID:94327554; PMID:8051087  
A;Accession: A53665  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-431 <DAV>  
A;Cross-references: GB:U10297; NID:g499266; PIDN:AAA57450.1; PID:g499267  
C;Superfamily: enolase

C;Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 42.6%; Score 52; DB 2; Length 431;  
Best Local Similarity 58.8%; Pred. No. 5.7;  
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALAICSQGAA 25

Db 107 ANAILGVSLAVCKAGAA 123

RESULT 10

A37210  
phosphopyruvate hydratase (EC 4.2.1.11) beta - rabbit  
N;Alternate names: enolase beta  
C;Species: Oryctolagus cuniculus (domestic rabbit)  
C;Date: 30-Dec-1991 #sequence\_revision 30-Dec-1991 #text\_change 05-Aug-1994  
C;Accession: A37210  
R;Chin, C.C.Q.  
J. Protein Chem. 9, 427-432, 1990  
A;Title: The primary structure of rabbit muscle enolase.  
A;Reference number: A37210; MUID:91113295; PMID:2275753  
A;Accession: A37210  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-433 <CHI>  
C;Superfamily: enolase  
C;Keywords: acetylated amino end; carbon-oxygen lyase; hydro-lyase  
F;1/Modified site: acetylated amino end (Ala) #status experimental

Query Match 42.6%; Score 52; DB 2; Length 433;  
Best Local Similarity 58.8%; Pred. No. 5.7;  
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALAICSQGAA 25

Db 107 ANAILGVSLAVCKAGAA 123

RESULT 11

NOHUG

phosphopyruvate hydratase (EC 4.2.1.11) gamma - human  
N;Alternate names: enolase gamma; neuron-specific enolase  
C;Species: Homo sapiens (man)  
C;Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 18-Jun-1999  
C;Accession: JU0060; S16163; S02077; I56569; S02616; S38303  
R;Oliva, D.; Barba, G.; Barbieri, G.; Giallongo, A.; Feo, S.  
Gene 79, 355-360, 1989  
A;Title: Cloning, expression and sequence homologies of cDNA for human gamma enolase  
A;Reference number: JU0060; MUID:90006764; PMID:2792767  
A;Accession: JU0060  
A;Molecule type: mRNA  
A;Residues: 1-434 <OL1>  
A;Cross-references: GB:M22349; NID:g951199; PIDN:AAB59554.1; PID:g1821116; GB:M27833  
R;Oliva, D.; Cali, L.; Feo, S.; Giallongo, A.  
Genomics 10, 157-165, 1991  
A;Title: Complete structure of the human gene encoding neuron-specific enolase.  
A;Reference number: S16163; MUID:91257823; PMID:2045099  
A;Accession: S16163  
A;Molecule type: DNA  
A;Residues: 1-434 <OL2>  
A;Cross-references: GB:X51956; NID:g31164; PIDN:CAA36215.1; PID:g31165  
R;McAleese, S.M.; Dunbar, B.; Fothergill, J.E.; Hinks, L.J.; Day, I.N.M.  
Eur. J. Biochem. 178, 413-417, 1988  
A;Title: Complete amino acid sequence of the neurone-specific gamma isozyme of enolase.  
A;Reference number: S02077; MUID:89091176; PMID:3208766  
A;Accession: S02077  
A;Molecule type: mRNA  
A;Residues: 2-3, Q', 5-239, M', 241-434 <MCA>  
A;Cross-references: EMBL:X13120; NID:g31145; PIDN:CAA31512.1; PID:g930063  
A;Note: part of this sequence was confirmed by protein sequencing  
A;Note: 264-Ala and 395-Ala were also found  
R;Van Obberghen, E.; Kamholz, J.; Bishop, J.G.



J. Neurosci. Res. 19, 450-456, 1988  
A;Title: Human gamma enolase: Isolation of a cDNA clone and expression in normal and tumor cells  
A;Reference number: 156569; MUID:88259288; PMID:3385803  
A;Accession: 156569  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 'GC', 29-126, 'N', 128-434 <VAN>  
A;Cross-references: GB:M36768; NID:g182117; PIDN:AAA52388.1; PID:g182118  
R;Day, I.N.M.; Allsopp, M.T.E.P.; Moore, D.C.M.; Thompson, R.J.  
FEBS Lett. 222, 139-143, 1987  
A;Title: Sequence conservation in the 3'-untranslated regions of neurone-specific enolase  
A;Reference number: S02616; MUID:88005129; PMID:3653393  
A;Accession: S02616  
A;Status: not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 425-434 <DAY>  
A;Cross-references: GB:Y00691; GB:M27610  
R;Harrington, C.R.; Quinn, G.B.; Hurt, J.; Day, I.N.M.; Wischik, C.M.  
Biochim. Biophys. Acta 1158, 120-128, 1993  
A;Title: Characterisation of an epitope specific to the neuron-specific isoform of human enolase  
A;Reference number: S38303; MUID:94002176; PMID:7691181  
A;Accession: S38303  
A;Molecule type: protein  
A;Residues: 156-173 <HAR>  
C;Comment: Enolase occurs with at least three isoforms (alpha, beta, and gamma) in mammalian cells. Thr-191 may be important for the enhanced tolerance to chloride ions of neuron-specific enolase.  
C;Genetics:  
A;Gene: GDB:ENO2  
A;Cross-references: GDB:119872; OMIM:131360  
A;Map position: 12p13-12p13  
A;Introns: 29/1; 61/1; 80/3; 104/1; 148/3; 223/1; 289/1; 356/2; 392/2; 412/2  
C;Complex: homodimer  
C;Function:  
A;Description: catalyzes the reversible dehydration of 2-phospho-D-glyceric acid to phosphoenolpyruvate  
A;Pathway: gluconeogenesis; glycolysis  
C;Superfamily: enolase  
C;Keywords: blocked amino end; brain; carbon-oxygen lyase; gluconeogenesis; glycolysis; F;2-434/Product: phosphopyruvate hydratase gamma #status predicted <MAT>  
F;2/Modified site: blocked amino end (Ser) (in mature form) (probably acetylated) #status predicted  
F;40/Binding site: magnesium 2 (Ser) #status predicted  
F;210,343/Active site: Glu, Lys #status predicted  
F;245,293,318/Binding site: magnesium 1 (Asp, Glu, Asp) #status predicted  
Query Match 42.6%; Score 52; DB 1; Length 434;  
Best Local Similarity 58.8%; Pred. No. 5.7;  
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
QY 9 ALALLGLALAICSQGAA 25  
Db 108 ANAILGVSLAVCKAGAA 124  
RESULT 12  
NOMSB  
phosphopyruvate hydratase (EC 4.2.1.11) beta - mouse  
N;Alternate names: enolase beta  
C;Species: Mus musculus (house mouse)  
C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 18-Jun-1999  
C;Accession: S17109; S18036; S29675; A33921  
R;Lamande, N.; Brosset, S.; Keller, A.; Lucas, M.; Lazar, M.  
submitted to the EMBL Data Library, September 1991  
A;Reference number: S17109  
A;Accession: S17109  
A;Molecule type: DNA  
A;Residues: 1-434 <LAM>  
A;Cross-references: EMBL:X61600; NID:g50848; PIDN:CAA43797.1; PID:g50849  
R;Peterson, C.A.; Cho, M.; Rastinejad, F.; Blau, H.M.  
submitted to the EMBL Data Library, October 1991  
A;Description: Beta-enolase: a gene expressed in undifferentiated postnatal myoblasts through the myoblast cell line C2C12.  
A;Reference number: S18036  
A;Accession: S18036  
A;Molecule type: mRNA

A;Residues: 1-434 <PET>  
A;Cross-references: EMBL:X62667; NID:g50143; PIDN:CAA44540.1; PID:g50144  
R;Lazar, M.; Lamande, N.; Brosset, S.; Lucas, M.; Keller, A.  
submitted to the EMBL Data Library, February 1991  
A;Reference number: S29675  
A;Accession: S29675  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-434 <LAZ>  
A;Cross-references: EMBL:X57747; NID:g50846; PIDN:CAA40913.1; PID:g50847  
R;Lamande, N.; Mazo, A.M.; Lucas, M.; Montarras, D.; Pinset, C.; Gros, F.; Legault-De-Proc. Natl. Acad. Sci. U.S.A. 86, 4445-4449, 1989  
A;Title: Murine muscle-specific enolase: cDNA cloning, sequence, and developmental expression  
A;Reference number: A33921; MUID:89282789; PMID:2734297  
A;Accession: A33921  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 59-233, 'NA', 236-434 <LA2>  
A;Cross-references: GB:M20745; NID:g193029; PIDN:AAA37554.1; PID:g387144  
A;Experimental source: skeletal muscle  
C;Genetics:  
A;Introns: 29/1; 61/1; 80/3; 104/1; 148/3; 223/1; 289/1; 356/2; 392/3; 412/2  
C;Function:  
A;Description: catalyzes the reversible dehydration of 2-phospho-D-glyceric acid to phosphoenolpyruvate  
A;Pathway: gluconeogenesis; glycolysis  
C;Superfamily: enolase  
C;Keywords: carbon-carbon lyase; carbon-oxygen lyase; dimer; gluconeogenesis; glycolysis  
F;40/Binding site: magnesium 2 (Ser) #status predicted  
F;210,343/Active site: Glu, Lys #status predicted  
F;245,293,318/Binding site: magnesium 1 (Asp, Glu, Asp) #status predicted  
Query Match 42.6%; Score 52; DB 1; Length 434;  
Best Local Similarity 58.8%; Pred. No. 5.7;  
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
QY 9 ALALLGLALAICSQGAA 25  
Db 108 ANAILGVSLAVCKAGAA 124  
RESULT 13  
A24742  
phosphopyruvate hydratase (EC 4.2.1.11) gamma - rat  
N;Alternate names: 2-phospho-glycerate dehydratase; enolase gamma; neuronal enolase  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 25-Oct-1987 #sequence\_revision 25-Oct-1987 #text\_change 22-Jun-1999  
C;Accession: A24742; PQ0006  
R;Sakimura, K.; Kushiya, E.; Obinata, M.; Odani, S.; Takahashi, Y.  
Proc. Natl. Acad. Sci. U.S.A. 82, 7453-7457, 1985  
A;Title: Molecular cloning and the nucleotide sequence of cDNA for neuron-specific enolase  
A;Reference number: A24742; MUID:86042683; PMID:2865729  
A;Accession: A24742  
A;Molecule type: mRNA  
A;Residues: 1-434 <SAK>  
A;Cross-references: GB:M11931; NID:g204041; PIDN:AAA41119.1; PID:g204042  
R;Sakimura, K.; Kushiya, E.; Takahashi, Y.; Suzuki, Y.  
Gene 60, 103-113, 1987  
A;Title: The structure and expression of neuron-specific enolase gene.  
A;Reference number: PQ0006; MUID:88152493; PMID:2450052  
A;Accession: PQ0006  
A;Molecule type: DNA  
A;Residues: 1-28 <SA2>  
A;Cross-references: GB:M22770; GB:M18742; NID:g205766; PIDN:AAA41725.1; PID:g554480  
C;Superfamily: enolase  
C;Keywords: carbon-oxygen lyase; gluconeogenesis; glycolysis; hydro-lyase  
Query Match 42.6%; Score 52; DB 2; Length 434;  
Best Local Similarity 58.8%; Pred. No. 5.7;  
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
QY 9 ALALLGLALAICSQGAA 25  
Db 108 ANAILGVSLAVCKAGAA 124



S10247

A32132

Search completed: March 21, 2003, 11:41:45  
Job time : 7.36481 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2003, 11:53:33 ; Search time 4.82833 Seconds  
(without alignments)  
276.816 Million cell updates/sec

Title: US-09-727-739B-19  
Perfect score: 122  
Sequence: 1 MRVSIHCALALLGLALAICSGAA 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 221153 seqs, 53462247 residues

Total number of hits satisfying chosen parameters: 221153

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID                   | Description       |
|------------|-------|-------------|--------|----------------------|-------------------|
| 1          | 50    | 41.0        | 1582   | 9 US-09-966-422B-11  | Sequence 11, Appl |
| 2          | 48    | 39.3        | 120    | 9 US-09-796-692-2277 | Sequence 2277, Ap |
| 3          | 48    | 39.3        | 162    | 9 US-09-796-692-2342 | Sequence 2342, Ap |
| 4          | 48    | 39.3        | 433    | 9 US-09-935-642-9    | Sequence 9, Appli |
| 5          | 48    | 39.3        | 1771   | 9 US-10-184-644-17   | Sequence 17, Appl |
| 6          | 47    | 38.5        | 423    | 10 US-09-843-164-10  | Sequence 10, Appl |
| 7          | 47    | 38.5        | 481    | 10 US-09-843-164-8   | Sequence 8, Appli |
| 8          | 47    | 38.5        | 502    | 10 US-09-843-164-14  | Sequence 14, Appl |
| 9          | 47    | 38.5        | 560    | 9 US-09-966-422B-2   | Sequence 2, Appli |
| 10         | 47    | 38.5        | 560    | 10 US-09-843-164-12  | Sequence 12, Appl |
| 11         | 47    | 38.5        | 769    | 10 US-09-843-164-2   | Sequence 2, Appli |
| 12         | 47    | 38.5        | 848    | 10 US-09-843-164-6   | Sequence 6, Appli |
| 13         | 46    | 37.7        | 400    | 9 US-09-971-228-12   | Sequence 12, Appl |
| 14         | 46    | 37.7        | 400    | 10 US-09-812-272-2   | Sequence 2, Appli |
| 15         | 46    | 37.7        | 400    | 12 US-10-037-616-27  | Sequence 27, Appl |
| 16         | 45    | 36.9        | 97     | 9 US-10-057-275-5    | Sequence 5, Appli |
| 17         | 45    | 36.9        | 97     | 9 US-09-920-137A-5   | Sequence 5, Appli |
| 18         | 45    | 36.9        | 419    | 10 US-09-789-919-44  | Sequence 44, Appl |
| 19         | 45    | 36.9        | 2694   | 9 US-10-184-644-207  | Sequence 207, App |

|    |    |      |     |   |                   |                   |
|----|----|------|-----|---|-------------------|-------------------|
| 20 | 44 | 36.1 | 239 | 9 | US-10-174-590-354 | Sequence 354, App |
| 21 | 44 | 36.1 | 239 | 9 | US-10-176-758-354 | Sequence 354, App |
| 22 | 44 | 36.1 | 239 | 9 | US-10-175-737-354 | Sequence 354, App |
| 23 | 44 | 36.1 | 239 | 9 | US-10-173-706-354 | Sequence 354, App |
| 24 | 44 | 36.1 | 239 | 9 | US-10-175-738-354 | Sequence 354, App |
| 25 | 44 | 36.1 | 239 | 9 | US-10-175-752-354 | Sequence 354, App |
| 26 | 44 | 36.1 | 239 | 9 | US-10-176-482-354 | Sequence 354, App |
| 27 | 44 | 36.1 | 239 | 9 | US-10-176-757-354 | Sequence 354, App |
| 28 | 44 | 36.1 | 239 | 9 | US-10-176-913-354 | Sequence 354, App |
| 29 | 44 | 36.1 | 239 | 9 | US-10-180-552-354 | Sequence 354, App |
| 30 | 44 | 36.1 | 239 | 9 | US-10-180-557-354 | Sequence 354, App |
| 31 | 44 | 36.1 | 239 | 9 | US-10-173-700-354 | Sequence 354, App |
| 32 | 44 | 36.1 | 239 | 9 | US-10-174-572-354 | Sequence 354, App |
| 33 | 44 | 36.1 | 239 | 9 | US-10-174-579-354 | Sequence 354, App |
| 34 | 44 | 36.1 | 239 | 9 | US-10-174-582-354 | Sequence 354, App |
| 35 | 44 | 36.1 | 239 | 9 | US-10-174-588-354 | Sequence 354, App |
| 36 | 44 | 36.1 | 239 | 9 | US-10-175-739-354 | Sequence 354, App |
| 37 | 44 | 36.1 | 239 | 9 | US-10-175-740-354 | Sequence 354, App |
| 38 | 44 | 36.1 | 239 | 9 | US-10-175-743-354 | Sequence 354, App |
| 39 | 44 | 36.1 | 239 | 9 | US-10-176-488-354 | Sequence 354, App |
| 40 | 44 | 36.1 | 239 | 9 | US-10-176-492-354 | Sequence 354, App |
| 41 | 44 | 36.1 | 239 | 9 | US-10-176-747-354 | Sequence 354, App |
| 42 | 44 | 36.1 | 239 | 9 | US-10-176-750-354 | Sequence 354, App |
| 43 | 44 | 36.1 | 239 | 9 | US-10-176-985-354 | Sequence 354, App |
| 44 | 44 | 36.1 | 239 | 9 | US-10-176-987-354 | Sequence 354, App |
| 45 | 44 | 36.1 | 239 | 9 | US-10-176-991-354 | Sequence 354, App |

ALIGNMENTS

RESULT 1

US-09-966-422B-11  
; Sequence 11, Application US/09966422B  
; Publication No. US20030044892A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBM6, EXPRESSE  
; FILE REFERENCE: D0040NP/3053-4119US3  
; CURRENT FILING DATE: 2002-05-07  
; PRIOR APPLICATION NUMBER: 60/235,602  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/306,604  
; PRIOR FILING DATE: 2001-07-19  
; PRIOR APPLICATION NUMBER: 60/315,412  
; PRIOR FILING DATE: 2001-08-28  
; NUMBER OF SEQ ID NOS: 81  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 11  
; LENGTH: 1582  
; TYPE: PRT  
; ORGANISM: Caenorhabditis elegans  
US-09-966-422B-11

Query Match 41.0%; Score 50; DB 9; Length 1582;  
Best Local Similarity 50.0%; Pred. No. 51;  
Matches 9; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 3 VSQIHCALALLGLALAIC 20  
||| |::: ||:::  
Db 910 VSTIGCAISIVCLALSVC 927

RESULT 2

US-09-796-692-2277  
; Sequence 2277, Application US/09796692  
; Publication No. US20020198362A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane



;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
;; FILE OF INVENTION: HEMATOLOGICAL MALIGNANCIES

;; FILE REFERENCE: 2077.001200  
;; CURRENT APPLICATION NUMBER: US/09/796,692

;; CURRENT FILING DATE: 2001-03-01  
;; PRIOR APPLICATION NUMBER: 60/186,126

;; PRIOR FILING DATE: 2000-03-01  
;; PRIOR APPLICATION NUMBER: 60/190,479

;; PRIOR FILING DATE: 2000-03-17  
;; PRIOR APPLICATION NUMBER: 60/200,545

;; PRIOR FILING DATE: 2000-04-27  
;; PRIOR APPLICATION NUMBER: 60/200,303

;; PRIOR FILING DATE: 2000-04-28  
;; PRIOR APPLICATION NUMBER: 60/200,779

;; PRIOR FILING DATE: 2000-04-28  
;; PRIOR APPLICATION NUMBER: 60/200,999

;; PRIOR FILING DATE: 2000-05-01  
;; PRIOR APPLICATION NUMBER: 60/202,084

;; PRIOR FILING DATE: 2000-05-04  
;; PRIOR APPLICATION NUMBER: 60/206,201

;; PRIOR FILING DATE: 2000-05-22  
;; PRIOR APPLICATION NUMBER: 60/218,950

;; PRIOR FILING DATE: 2000-07-14  
;; PRIOR APPLICATION NUMBER: 60/222,903

;; PRIOR FILING DATE: 2000-08-03  
;; PRIOR APPLICATION NUMBER: 60/223,416

;; PRIOR FILING DATE: 2000-08-04  
;; PRIOR APPLICATION NUMBER: 60/223,378

;; PRIOR FILING DATE: 2000-08-07  
;; NUMBER OF SEQ ID NOS: 9597

;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 2277

;; LENGTH: 120  
;; TYPE: PRT

;; ORGANISM: Homo sapiens  
;; FEATURE:

;; NAME/KEY: variant  
;; LOCATION: (1)...(120)

;; OTHER INFORMATION: Xaa - Any amino acid  
US-09-796-692-2277

QY 9 ALALLGLALAICSQGA 24  
| : : : : : | |

Db 65 ANAILGVSLAVCKAGA 80

RESULT 3

-09-796-692-2342  
Sequence 2342, Application US/09796692

;; Publication No. US20020198362A1  
;; GENERAL INFORMATION:

;; APPLICANT: Galger, Alexander  
;; APPLICANT: Mannion, Jane

;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
;; FILE REFERENCE: 2077.001200

;; CURRENT APPLICATION NUMBER: US/09/796,692  
;; CURRENT FILING DATE: 2001-03-01

;; PRIOR APPLICATION NUMBER: 60/186,126  
;; PRIOR FILING DATE: 2000-03-01

;; PRIOR APPLICATION NUMBER: 60/190,479  
;; PRIOR FILING DATE: 2000-03-17

;; PRIOR APPLICATION NUMBER: 60/200,545  
;; PRIOR FILING DATE: 2000-04-27

;; PRIOR APPLICATION NUMBER: 60/200,303  
;; PRIOR FILING DATE: 2000-04-28

;; PRIOR APPLICATION NUMBER: 60/200,779  
;; PRIOR FILING DATE: 2000-04-28

;; PRIOR APPLICATION NUMBER: 60/200,999  
;; PRIOR FILING DATE: 2000-05-01

;; PRIOR APPLICATION NUMBER: 60/202,084  
;; PRIOR FILING DATE: 2000-05-04

;; PRIOR APPLICATION NUMBER: 60/206,201  
;; PRIOR FILING DATE: 2000-05-22

;; PRIOR APPLICATION NUMBER: 60/218,950  
;; PRIOR FILING DATE: 2000-07-14

;; PRIOR APPLICATION NUMBER: 60/222,903  
;; PRIOR FILING DATE: 2000-08-03

;; PRIOR APPLICATION NUMBER: 60/223,416  
;; PRIOR FILING DATE: 2000-08-04

;; PRIOR APPLICATION NUMBER: 60/223,378  
;; PRIOR FILING DATE: 2000-08-07

;; NUMBER OF SEQ ID NOS: 9597  
;; SOFTWARE: FastSeq for Windows Version 3.0

;; SEQ ID NO 2342  
;; LENGTH: 162

;; TYPE: PRT  
;; ORGANISM: Homo sapiens

;; FEATURE:  
;; NAME/KEY: variant

;; LOCATION: (1)...(162)  
;; OTHER INFORMATION: Xaa - Any amino acid

US-09-796-692-2342

Query Match 39.3%; Score 48; DB 9; Length 162;  
Best Local Similarity 56.2%; Pred. No. 11;

Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALAICSQGA 24  
| : : : : : | |

Db 109 ANAILGVSLAVCKAGA 124

RESULT 4

US-09-935-642-9  
Sequence 9, Application US/09935642

;; Publication No. US20030044795A1  
;; GENERAL INFORMATION:

;; APPLICANT: BYRJALSEN, Inger  
;; APPLICANT: LARSEN, Peter

;; TITLE OF INVENTION: Biochemical Markers for the Human  
;; FILE REFERENCE: 8969-014

;; CURRENT APPLICATION NUMBER: US/09/935,642  
;; CURRENT FILING DATE: 2001-08-24

;; PRIOR APPLICATION NUMBER: PCT/GB97/02394  
;; PRIOR FILING DATE: 1997-09-05

;; PRIOR APPLICATION NUMBER: PCT/GB9707132.8  
;; PRIOR FILING DATE: 1997-04-08

;; PRIOR APPLICATION NUMBER: PCT/GB9618600.2  
;; PRIOR FILING DATE: 1996-09-06

;; NUMBER OF SEQ ID NOS: 16  
;; SOFTWARE: FastSeq for Windows Version 4.0

;; SEQ ID NO 9  
;; LENGTH: 433

;; TYPE: PRT  
;; ORGANISM: Homo sapiens

US-09-935-642-9

Query Match 39.3%; Score 48; DB 9; Length 433;  
Best Local Similarity 56.2%; Pred. No. 27;

Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALAICSQGA 24  
| : : : : : | |

Db 107 ANAILGVSLAVCKAGA 122

RESULT 5

US-10-184-644-17



```

; Sequence 17, Application US/10184644
; Publication NO. US20030044930A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TR
; TITLE OF INVENTION: ACIDS ENCODING
; FILE REFERENCE: P3430R1C227
; CURRENT APPLICATION NUMBER: US/10/1
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See Fil
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 17
; -LENGTH: 1771
; TYPE: DNA
; ORGANISM: Homo Sapien
; US-10-184-644-17

```

```

Query Match      39.3%;   Score 48;   DB 9;   Length 1771;
Best Local Similarity 55.6%;
Pred. NO. 1.1e+02;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

```

RESULT 6  
US-09-843-164-10  
; Sequence 10, Application US/09843164  
; Patent No. US20020061556A1  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Wang, Xiaoming  
; APPLICANT: Scoville, John  
; TITLE OF INVENTION: No. US20020061556A1e Human Membrane Proteins and Polynucleotides  
; FILE REFERENCE: 07705.0014-00000  
; CURRENT APPLICATION NUMBER: US/09/843,164  
; CURRENT FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: US 60/199,950  
; PRIOR FILING DATE: 2000-04-27  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 423  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-09-843-164-10

RESULT 7  
US-09-843-164-8  
; Sequence 8, Application US/09843164  
; Patent No. US2002061556A1  
; GENERAL INFORMATION:  
; APPLICANT: walke, D. wade

```

; APPLICANT: Wang, Xiaoming
; APPLICANT: Scoville, John
; TITLE OF INVENTION: NO. US20020061556A1 Human Membrane Proteins and Polynucleotides
; FILE REFERENCE: 07705.0014-00000
; CURRENT APPLICATION NUMBER: US/09/843,164
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/199,950
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 8
; LENGTH: 481
; TYPE: PRT
; ORGANISM: homo sapiens
; US-09-843-164-8

```

RESULT 8  
US-09-843-164-14  
; Sequence 14, Application US/09843164  
; Patent No. US20020061556A1  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Wang, Xiaoming  
; APPLICANT: Scoville, John  
; TITLE OF INVENTION: No. US20020061556A1el Human Membrane Proteins and Polynucle  
; FILE REFERENCE: 07705.0014-00000  
; CURRENT APPLICATION NUMBER: US/09/843;164  
; CURRENT FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: US 60/199,950  
; PRIOR FILING DATE: 2000-04-27  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 14  
; LENGTH: 502  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-09-843-164-14

RESULT 9  
US-09-966-422B-2  
; Sequence 2, Application US/09966422B  
; Publication No. US20030044892A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBMY6, EXPRESSED  
; TITLE OF INVENTION: SMALL INTESTINE  
; FILE REFERENCE: D0040NP/3053-4119US3  
; CURRENT APPLICATION NUMBER: US/09/966,422B  
; CURRENT FILING DATE: 2002-05-07  
; PRIOR APPLICATION NUMBER: 60/235,602  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/306,604  
; PRIOR FILING DATE: 2001-07-19  
; PRIOR APPLICATION NUMBER: 60/315,412  
; PRIOR FILING DATE: 2001-08-28



; NUMBER OF SEQ ID NOS: 81  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2  
; LENGTH: 560  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-966-422B-2

Query Match 38.5%; Score 47; DB 9; Length 560;  
Best Local Similarity 45.0%; Pred. No. 48;  
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 3 VSQIHICALALLGLALAICSQ 22  
; : | : | | : | | | : |  
Db 200 LSNVGCALSVTGLALTIVIFQ 219

RESULT 10

US-09-843-164-12  
; Sequence 12, Application US/09843164  
; Patent No. US20020061556A1

; GENERAL INFORMATION:

; APPLICANT: Walke, D. Wade

; APPLICANT: Wang, Xiaoming

; APPLICANT: Scoville, John

; TITLE OF INVENTION: No. US20020061556A1el Human Membrane Proteins and Polynucleotides

; FILE REFERENCE: 07705.0014-00000

; CURRENT APPLICATION NUMBER: US/09/843,164

; CURRENT FILING DATE: 2001-04-27

; PRIOR FILING DATE: 2000-04-27

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 12

; LENGTH: 560

; TYPE: PRT

; ORGANISM: homo sapiens

US-09-843-164-12

Query Match 38.5%; Score 47; DB 10; Length 560;  
Best Local Similarity 45.0%; Pred. No. 48;  
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 3 VSQIHICALALLGLALAICSQ 22  
; : | : | | : | | | : |  
Db 200 LSNVGCALSVTGLALTIVIFQ 219

RESULT 11

US-09-843-164-2  
; Sequence 2, Application US/09843164  
; Patent No. US20020061556A1

; GENERAL INFORMATION:

; APPLICANT: Walke, D. Wade

; APPLICANT: Wang, Xiaoming

; APPLICANT: Scoville, John

; TITLE OF INVENTION: No. US20020061556A1el Human Membrane Proteins and Polynucleotides

; FILE REFERENCE: 07705.0014-00000

; CURRENT APPLICATION NUMBER: US/09/843,164

; CURRENT FILING DATE: 2001-04-27

; PRIOR FILING DATE: 2000-04-27

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 769

; TYPE: PRT

; ORGANISM: homo sapiens

US-09-843-164-2

Query Match 38.5%; Score 47; DB 10; Length 769;  
Best Local Similarity 45.0%; Pred. No. 66;  
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;



Db 195 LFCVLAFLGLAICA 210

RESULT 14

US-09-812-272-2  
; Sequence 2, Application US/09812272  
; Patent No. US20020045214A1  
; GENERAL INFORMATION:  
; APPLICANT: Pamela Lane  
; APPLICANT: Ping Tsui  
; APPLICANT: Nabil A. Elshourbagy  
; TITLE OF INVENTION: RAT G-PROTEIN COUPLED RECEPTOR AXOR29  
; FILE REFERENCE: GP-70685  
; CURRENT APPLICATION NUMBER: US/09/812,272  
; CURRENT FILING DATE: 2001-03-20  
; PRIOR APPLICATION NUMBER: 60/191,153  
; PRIOR FILING DATE: 2000-03-22  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: FastSeq for Windows version 3.0  
; SEQ ID NO 2  
; LENGTH: 400  
; TYPE: PRT  
; ORGANISM: HOMO SAPIENS  
US-09-812-272-2

Query Match 37.7%; Score 46; DB 10; Length 400;  
Best Local Similarity 50.0%; Pred. No. 48;  
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 6 IHCALALGLAICA 21  
: | | | | :  
Db 195 LFCVLAFLGLAICA 210

RESULT 15

US-10-037-616-27  
; Sequence 27, Application US/10037616  
; Patent No. US20020123148A1  
; GENERAL INFORMATION:  
; APPLICANT: English, Denis  
; APPLICANT: Kovacs, Richard J.  
; APPLICANT: Rizzo, Maria T.  
; APPLICANT: Sliva, Daniel T.  
; TITLE OF INVENTION: Sphingolipid Compositions and Methods for Their Therapeutic Use  
; FILE REFERENCE: 7042-119  
; CURRENT APPLICATION NUMBER: US/10/037,616  
; CURRENT FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: US 60/243,887  
; PRIOR FILING DATE: 2000-10-27  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 27  
; LENGTH: 400  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-10-037-616-27

Query Match 37.7%; Score 46; DB 12; Length 400;  
Best Local Similarity 50.0%; Pred. No. 48;  
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 6 IHCALALGLAICA 21  
: | | | | :  
Db 195 LFCVLAFLGLAICA 210

Search completed: March 21, 2003, 11:56:32  
Job time : 5.82833 secs







CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Patrea L. Pabst  
 STREET: 2800 One Atlantic Center, 1201 W. Peachtree  
 STREET: St.  
 CITY: Atlanta  
 STATE: GA  
 COUNTRY: USA  
 ZIP: 30309-3450  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/974,691  
 FILING DATE: 20-NOV-1997

CLASSIFICATION: /  
PRIOR APPLICATION DATA: / US 60/031,196 /  
APPLICATION NUMBER: / 20-NOV-1996 /  
FILING DATE: /  
PRIOR APPLICATION DATA: /  
APPLICATION NUMBER: / US 60/046,126 /  
FILING DATE: / 09-MAY-1997 /  
ATTORNEY/AGENT INFORMATION: /  
NAME: Pabst, Patrea L. /  
REGISTRATION NUMBER: 31,284 /  
REFERENCE/DOCKET NUMBER: OMRF 166 /  
TELECOMMUNICATION INFORMATION: /  
TELEPHONE: 404-873-8794 /  
TELEFAX: 404-873-8795 /  
INFORMATION FOR SEQ ID NO: 6: /  
SEQUENCE CHARACTERISTICS: /  
LENGTH: 445 amino acids /  
TYPE: amino acid /  
STRANDEDNESS: single /  
TOPOLOGY: linear /  
MOLECULE TYPE: protein /  
ORIGINAL SOURCE: /  
ORGANISM: Homo sapiens /  
US-08-974-691-6 /

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/974,691  
 FILING DATE: 20-NOV-1997  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/031,196  
 FILING DATE: 20-NOV-1996  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/046,126  
 FILING DATE: 09-MAY-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Pabst, Patrea L.  
 REGISTRATION NUMBER: 31,284  
 REFERENCE/DOCKET NUMBER: OMRF 166  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 404-873-8794  
 TELEFAX: 404-873-8795  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 451 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 ORIGINAL SOURCE:  
 ORGANISM: Homo sapiens  
 US-08-974-691-2

|                       |                 |                            |        |             |
|-----------------------|-----------------|----------------------------|--------|-------------|
| Query Match           | 37.7%           | Score 46;                  | DB 4;  | Length 451; |
| Best Local Similarity | 42.9%           | Pred. NO. 23;              |        |             |
| Matches 9;            | Conservative 3; | Mismatches 9;              | Indels |             |
| QY                    | 5               | QIHCA LALLGLALAI CSQGA A   | 25     |             |
|                       |                 |                            |        |             |
| Db                    | 259             | QIHM ERV KVGSR LTL CAOG CA | 279    |             |

RESULT 4  
US-08-974-691-3  
Sequence 3, Application US/08974691  
Patent No. 6225103  
GENERAL INFORMATION:  
APPLICANT: Keolsch, Gerald  
APPLICANT: Lin, Xinli  
APPLICANT: Tang, Jordan  
TITLE OF INVENTION: Cloning and Characterization of Napsin  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Patrea L. Pabst  
STREET: 2800 One Atlantic Center, 1201 W. Peachtree  
STREET: St.  
CITY: Atlanta  
STATE: GA  
COUNTRY: USA  
ZIP: 30309-3450  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/974,691  
FILING DATE: 20-NOV-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/031,196  
FILING DATE: 20-NOV-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/046,126  
FILING DATE: 09-MAY-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284



REFERENCE/DOCKET NUMBER: OMRF 166  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404-873-8794  
TELEFAX: 404-873-8795  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 419 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-974-691-3

Query Match 36.9%; Score 45; DB 4; Length 419;  
Best Local Similarity 33.3%; Pred. No. 30;  
Matches 7; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 5 QIHCALALLGLALAIQSGAA 25  
Db 254 QVHVESVKVGTGLSLCAQGS 274

RESULT 5  
US-09-705-448-10  
Sequence 10, Application US/09705448  
Patent No. 6432690  
GENERAL INFORMATION:  
APPLICANT: Xu, Hong  
APPLICANT: Bruno, Sandra A.  
APPLICANT: Elsenboss, Laura A.  
APPLICANT: Fogliano, Michael  
APPLICANT: Cohan, Victoria L.  
APPLICANT: Bandman, Olga  
TITLE OF INVENTION: HUMAN ASPARTIC PROTEASES  
FILE REFERENCE: PF-0458-1 CIP  
CURRENT APPLICATION NUMBER: US/09/705,448  
CURRENT FILING DATE: 2000-11-02  
PRIOR APPLICATION NUMBER: 09/116,641  
PRIOR FILING DATE: 1998-07-16  
PRIOR APPLICATION NUMBER: 09/008,271  
PRIOR FILING DATE: 1998-01-16  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 10  
LENGTH: 419  
TYPE: PRT  
ORGANISM: MUS MUSCULUS  
FEATURE:  
OTHER INFORMATION: 1906810, GenBank  
US-09-705-448-10

Query Match 36.9%; Score 45; DB 4; Length 419;  
Best Local Similarity 33.3%; Pred. No. 30;  
Matches 7; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 5 QIHCALALLGLALAIQSGAA 25  
Db 254 QVHVESVKVGTGLSLCAQGS 274

RESULT 6  
US-08-637-670-25  
Sequence 25, Application US/08637670  
Patent No. 6413521  
GENERAL INFORMATION:  
APPLICANT: MCMICHAEL-PHILLIPS et al.  
TITLE OF INVENTION: Helminth Parasite Antigen with  
TITLE OF INVENTION: Aminopeptidase-like Activity  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Barbara G. Ernst  
STREET: 555 13TH STREET, NW Suite 701E  
CITY: Washington

STATE: DC  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/637,670  
FILING DATE: 26-JUN-1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Ernst, Barbara G.  
REGISTRATION NUMBER: 30,377  
REFERENCE/DOCKET NUMBER: 1811-0232  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-783-6040  
TELEFAX: 202-783-6031  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 350 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-637-670-25

Query Match 36.1%; Score 44; DB 4; Length 350;  
Best Local Similarity 42.1%; Pred. No. 34;  
Matches 8; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 1 MRVSOIHCALALLGLALAI 19  
Db 12 LRLTPITLSIALGLIAYAV 30

RESULT 7  
US-09-105-343A-2  
Sequence 2, Application US/09105343A  
Patent No. 6207642  
GENERAL INFORMATION:  
APPLICANT: WILEY, S.R.  
TITLE OF INVENTION: MEMBER OF THE TNF FAMILY USEFUL  
TITLE OF INVENTION: FOR TREATMENT AND DIAGNOSIS OF DISEASE  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Abbott Laboratories  
STREET: 100 Abbott Park Road  
CITY: Abbott Park  
STATE: IL  
COUNTRY: USA  
ZIP: 60064-6050  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: FastSeq for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,343A  
FILING DATE: 12-FEB-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US98/02859  
FILING DATE: 12-FEB-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: BECKER, Cheryl L.  
REGISTRATION NUMBER: 35,441  
REFERENCE/DOCKET NUMBER: 6048.US.P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 847-935-1729  
TELEFAX: 847-938-2623  
TELEX:



INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 249 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: No. 6207642e  
US-09-105-343A-2

Query Match 35.2%; Score 43; DB 4; Length 249;  
Best Local Similarity 62.5%; Pred. No. 33;  
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 9 ALALLGLALAICSGQA 24  
Db 31 ALACGLLLAVVSLGS 46

RESULT 8  
US-08-469-667-14  
Sequence 14, Application US/08469667  
Patent No. 5733748  
GENERAL INFORMATION:  
APPLICANT: Yu, Guo-Liang  
TITLE OF INVENTION: Colon Specific Genes and Proteins  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,  
ADDRESSEE: Stewart & Olstein  
STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: NJ  
COUNTRY: USA  
ZIP: 07068-1739  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION NUMBER: US/08/469,667  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Ferraro, Gregory D.  
REGISTRATION NUMBER: 36,134  
REFERENCE/DOCKET NUMBER: 325800-435  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 81 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-469-667-14

Query Match 34.4%; Score 42; DB 1; Length 81;  
Best Local Similarity 52.9%; Pred. No. 13;  
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 9 ALALLGLALAICSGQA 25  
Db 5 ALCMLGLVLLSSSSA 21

RESULT 9  
US-09-224-110-14  
Sequence 14, Application US/09224110  
Patent No. 6337195  
GENERAL INFORMATION:

APPLICANT: Yu, Guo-Liang  
APPLICANT: Rosen, Craig  
TITLE OF INVENTION: Colon Specific Genes and Proteins  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,  
ADDRESSEE: Stewart & Olstein  
STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: NJ  
COUNTRY: USA  
ZIP: 07068-1739  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION NUMBER: US/09/224,110  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/469,667  
FILING DATE: 06-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Ferraro, Gregory D.  
REGISTRATION NUMBER: 36,134  
REFERENCE/DOCKET NUMBER: 325800-435  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 81 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-224-110-14

Query Match 34.4%; Score 42; DB 4; Length 81;  
Best Local Similarity 52.9%; Pred. No. 13;  
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 9 ALALLGLALAICSGQA 25  
Db 5 ALCMLGLVLLSSSSA 21

RESULT 10  
PCT-US95-07289-14  
Sequence 14, Application PC/TUS9507289  
GENERAL INFORMATION:  
APPLICANT: Yu, Guo-Liang  
APPLICANT: Rosen, Craig  
TITLE OF INVENTION: Colon Specific Genes and Proteins  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,  
ADDRESSEE: Stewart & Olstein  
STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: NJ  
COUNTRY: USA  
ZIP: 07068-1739  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION NUMBER: PCT/US95/07289  
FILING DATE: 06-JUN-1995  
CLASSIFICATION:



ATTORNEY/AGENT INFORMATION:  
NAME: Ferraro, Gregory D.  
REGISTRATION NUMBER: 36,134  
REFERENCE/DOCKET NUMBER: 325800-265  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 81 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-07289-14

Query Match 34.4%; Score 42; DB 5; Length 81;  
Best Local Similarity 52.9%; Pred. No. 13;  
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 9 ALALLGLALICSGAA 25  
|| :||| ||: | :|  
Db 5 ALCMLGLVLLSSSA 21

RESULT 11  
US-08-723-938-3  
Sequence 3, Application US/08723938  
Patent No. 5776759  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Coleman, Roger  
TITLE OF INVENTION: TWO NOVEL HUMAN CATHESPIN PROTEINS  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: U.S.  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/723,938  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0125 US  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 395 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
IMMEDIATE SOURCE:  
LIBRARY: LUNGNOT02

CLONE: 312099  
US-08-723-938-3  
Query Match 34.4%; Score 42; DB 1; Length 395;  
Best Local Similarity 38.1%; Pred. No. 80;  
Matches 8; Conservative 4; Mismatches 9; Indels 0; Gaps 0;  
QY 5 QIHCALALLGLALICSGAA 25  
||| :| | :| :|  
Db 259 QIHMERVKVGPGLTLCAGCA 279

RESULT 12  
US-09-080-538-3  
Sequence 3, Application US/09080538  
Patent No. 5965129  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Coleman, Roger  
TITLE OF INVENTION: TWO NOVEL HUMAN CATHESPIN PROTEINS  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: U.S.  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/080,538  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/723,938  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0125 US  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 395 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
IMMEDIATE SOURCE:  
LIBRARY: LUNGNOT02  
CLONE: 312099  
US-09-080-538-3

Query Match 34.4%; Score 42; DB 2; Length 395;  
Best Local Similarity 38.1%; Pred. No. 80;  
Matches 8; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 5 QIHCALALLGLALICSGAA 25  
||| :| | :| :|  
Db 259 QIHMERVKVGPGLTLCAGCA 279



## RESULT 13

US-09-008-271A-4  
; Sequence 4, Application US/09008271A  
; Patent No. 6203979

## GENERAL INFORMATION:

APPLICANT: Bandman, Olga  
Hillman, Jennifer L.  
Yue, Henry  
Guegler, Karl J.  
Corley, Neil C.  
Tang, Tom Y.  
Shah, Purvi

TITLE OF INVENTION: HUMAN PROTEASE MOLECULES

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Dr.

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/008,271A

FILING DATE: 16-Jan-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: <Unknown>

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Mohan-Peterson, Sheela

REGISTRATION NUMBER: 41,201

REFERENCE/DOCKET NUMBER: PF-0458 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 420 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: LUNGAST01

CLONE: 877617

SEQUENCE DESCRIPTION: SEQ ID NO: 4 :

US-09-008-271A-4

Query Match 34.4%; Score 42; DB 4; Length 420;

Best Local Similarity 38.1%; Pred. No. 85;

Matches 8; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 5 QIHCAALGLALALICSGAA 25

||| :| :| :| :|

Db 259 QIHMERVKVGPGLTLCAGCA 279

## RESULT 14

US-08-974-691-8

; Sequence 8, Application US/08974691

; Patent No. 6225103

; GENERAL INFORMATION:

APPLICANT: Keolsch, Gerald

APPLICANT: Lin, Xinli

APPLICANT: Tang, Jordan

TITLE OF INVENTION: Cloning and Characterization of Napsin

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Patrea L. Pabst

STREET: 2800 One Atlantic Center, 1201 W. Peachtree

STREET: St.  
CITY: Atlanta  
STATE: GA  
COUNTRY: USA  
ZIP: 30309-3450  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/974,691  
FILING DATE: 20-NOV-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/031,196  
FILING DATE: 20-NOV-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/046,126  
FILING DATE: 09-MAY-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: OMRP 166  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404-873-8794  
TELEFAX: 404-873-8795  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 420 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
US-08-974-691-8

Query Match 34.4%; Score 42; DB 4; Length 420;

Best Local Similarity 38.1%; Pred. No. 85;

Matches 8; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 5 QIHCAALGLALALICSGAA 25

||| :| :| :| :|

Db 259 QIHMERVKVGPGLTLCAGCA 279

## RESULT 15

US-09-705-448-1

; Sequence 1, Application US/09705448

; Patent No. 6432690

; GENERAL INFORMATION:

APPLICANT: Xu, Hong

APPLICANT: Bruno, Sandra A.

APPLICANT: Eisenboss, Laura A.

APPLICANT: Fogliano, Michael

APPLICANT: Cohan, Victoria L.

APPLICANT: Bandman, Olga

TITLE OF INVENTION: HUMAN ASPARTIC PROTEASES

FILE REFERENCE: PF-0458-1 CIP

CURRENT APPLICATION NUMBER: US/09/705,448

CURRENT FILING DATE: 2000-11-02

PRIOR APPLICATION NUMBER: 09/116,641

PRIOR FILING DATE: 1998-07-16

PRIOR APPLICATION NUMBER: 09/008,271

PRIOR FILING DATE: 1998-01-16

NUMBER OF SEQ ID NOS: 10

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 1

LENGTH: 420

TYPE: PRT

ORGANISM: HOMO SAPIENS

FEATURE:



us-09-727-739b-19.ra1

Page 7.

Query Match 34.4%; Score 42; DB 4; Length 420;  
Best Local Similarity 38.1%; Pred. No: 85;  
Matches 8; Conservative 4; Mismatches 9; Indels

|     |         |                       |              |    |            |    |        |    |      |    |
|-----|---------|-----------------------|--------------|----|------------|----|--------|----|------|----|
|     | Matches | 8;                    | Conservative | 4; | Mismatches | 9; | Indels | 0; | Gaps | 0; |
| QY: | 5       | QHICALALLGLALAICSQGAA | 25           |    |            |    |        |    |      |    |
|     |         |                       | :            | :  | :          | :  | :      | :  | :    | :  |
|     |         |                       |              |    |            |    |        |    |      |    |
| Dd  | 259     | QIHMERVKVGPGLTLCAKGC  | 279          |    |            |    |        |    |      |    |

Search completed: March 21, 2003, 11:42:36  
Job time : 5.72103 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2003, 14:35:23 ; Search time 15.4506 seconds  
(without alignments)  
215.607 Million cell updates/sec

Title: US-09-727-739B-19  
Perfect score: 122  
Sequence: 1 MRVSIHCALALLGLALAICSQGAA 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues.

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_101002:\*

|     |                                                   |
|-----|---------------------------------------------------|
| 1:  | /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:* |
| 2:  | /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:* |
| 3:  | /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:* |
| 4:  | /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:* |
| 5:  | /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:* |
| 6:  | /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:* |
| 7:  | /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:* |
| 8:  | /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:* |
| 9:  | /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:* |
| 10: | /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:* |
| 11: | /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:* |
| 12: | /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:* |
| 13: | /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:* |
| 14: | /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:* |
| 15: | /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:* |
| 16: | /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:* |
| 17: | /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:* |
| 18: | /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:* |
| 19: | /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:* |
| 20: | /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:* |
| 21: | /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:* |
| 22: | /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:* |
| 23: | /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:* |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID       | Description         |
|------------|-------|-------------|--------|-------------|---------------------|
| 1          | 122   | 100.0       | 111    | 22 AAU07668 | Rainbow trout prep  |
| 2          | 110   | 90.2        | 115    | 22 AAU07667 | Rainbow trout prep  |
| 3          | 64    | 52.5        | 114    | 22 AAU07666 | Rainbow trout prep  |
| 4          | 52    | 42.6        | 433    | 18 AAW14001 | Enolase protein.    |
| 5          | 52    | 42.6        | 434    | 23 AAM48922 | Human enolase 3 (be |
| 6          | 48    | 39.3        | 120    | 22 AAM81913 | Human haematologic  |
| 7          | 48    | 39.3        | 162    | 22 AAM81978 | Human haematologic  |
| 8          | 48    | 39.3        | 429    | 21 AAB42064 | Human ORF1828       |
| 9          | 48    | 39.3        | 433    | 19 AAW54357 | Alpha Enolase.. Ho  |
| 10         | 48    | 39.3        | 434    | 23 ABB57379 | Rat mucocardial ce  |

|    |      |      |      |    |          |                    |
|----|------|------|------|----|----------|--------------------|
| 11 | 47   | 38.5 | 210  | 20 | AAU35223 | Chlamydia pneumoni |
| 12 | 47   | 38.5 | 423  | 23 | AAU47765 | Novel G-protein co |
| 13 | 47   | 38.5 | 472  | 23 | AAU93172 | Arabidopsis transc |
| 14 | 47   | 38.5 | 475  | 23 | AAU93018 | Arabidopsis transc |
| 15 | 47   | 38.5 | 481  | 23 | AAU47764 | Novel G-protein co |
| 16 | 47   | 38.5 | 502  | 23 | AAU47767 | Novel G-protein co |
| 17 | 47   | 38.5 | 560  | 23 | ABG34861 | Human G-protein co |
| 18 | 47   | 38.5 | 560  | 23 | AAU47766 | Novel G-protein co |
| 19 | 47   | 38.5 | 674  | 22 | AAU47766 | Human protein sequ |
| 20 | 47   | 38.5 | 682  | 23 | AAO15109 | Human G protein-co |
| 21 | 47   | 38.5 | 718  | 22 | AAU92973 | Human protein sequ |
| 22 | 47   | 38.5 | 724  | 22 | ABG17176 | Novel human diagno |
| 23 | 47   | 38.5 | 769  | 23 | AAU47761 | Novel G-protein co |
| 24 | 47   | 38.5 | 797  | 22 | AAU95679 | Human protein sequ |
| 25 | 47   | 38.5 | 797  | 23 | AAE23412 | Human G-protein co |
| 26 | 47   | 38.5 | 798  | 22 | AAU93238 | Human protein sequ |
| 27 | 47   | 38.5 | 848  | 23 | AAU47763 | Novel G-protein co |
| 28 | 47   | 38.5 | 1150 | 22 | AAM40294 | Human polypeptide  |
| 29 | 46   | 37.7 | 400  | 22 | AAU72561 | Rat G protein-coup |
| 30 | 46   | 37.7 | 400  | 23 | AAO14445 | Rattus norvegicus  |
| 31 | 46   | 37.7 | 433  | 21 | AAU44810 | Human Aspartic Pro |
| 32 | 46   | 37.7 | 438  | 19 | AAW57042 | Human aspartic pro |
| 33 | 46   | 37.7 | 451  | 19 | AAW54877 | Human napsin A pro |
| 34 | 46   | 37.7 | 1982 | 22 | ABG09731 | Novel human diagno |
| 35 | 45   | 36.9 | 311  | 21 | AAU44937 | Corn sulphate perm |
| 36 | 45   | 36.9 | 395  | 21 | AAG14553 | Arabidopsis thalia |
| 37 | 45   | 36.9 | 395  | 21 | AAG52413 | Arabidopsis thalia |
| 38 | 45   | 36.9 | 419  | 21 | AAU79177 | Haematopoietic ste |
| 39 | 45   | 36.9 | 425  | 21 | AAG14552 | Arabidopsis thalia |
| 40 | 45   | 36.9 | 425  | 21 | AAG52412 | Arabidopsis thalia |
| 41 | 45   | 36.9 | 453  | 21 | AAG14551 | Arabidopsis thalia |
| 42 | 45   | 36.9 | 453  | 21 | AAG52411 | Arabidopsis thalia |
| 43 | 45   | 36.9 | 453  | 23 | AAU93186 | Arabidopsis thalia |
| 44 | 45   | 36.9 | 1039 | 22 | ABB63666 | Arabidopsis transc |
| 45 | 44.5 | 36.5 | 116  | 21 | AAG03774 | Drosophila melanog |
|    |      |      |      |    |          | Human secreted pro |

ALIGNMENTS

RESULT 1  
AAU07668  
ID AAU07668 standard; Protein; 111 AA.

XX AAU07668;

AC AAU07668;

DT 04-DEC-2001 (first entry)

XX Rainbow trout preprosomatostatin II (PPSS-II'') polypeptide.

XX Rainbow trout; somatostatin; preprosomatostatin; hypersecretion; PPSS-I;  
PPSS-II'; PPSS-II'; endocrine tumour; pituitary gland; glucagonoma; AIDS;  
KW gastroenteropancreatic tissue; acromegaly; gastrinoma; diabetes mellitus;  
KW carcinoid syndrome; cell proliferation; apoptosis; growth hormone;  
KW glucagon; acquired immunodeficiency syndrome; neurological disorder; HIV;  
KW epilepsy; Alzheimer's disease; Huntington's disease; neuroprotective;  
KW neoplasm; metastasis; gene therapy; antidiabetic; nootropic; cytostatic;  
KW anti-human immunodeficiency virus; osteopathic; anticonvulsant.

XX Oncorhynchus mykiss.

OS Key Location/Qualifiers

FT Peptide 1..25

FT Protein /note= "Signal peptide"

FT Protein /note= "ppss-II'' pre-sequence"

FT Protein /note= "Mature PPSS-II''"

FT Peptide /note= "PPSS-II'' pro-sequence"

FT Peptide /note= "Prosomatostatin II''"

FT Cleavage-site 96..97



FT Peptide /note= "Dibasic cleavage site"  
FT 98..111  
FT /note= "SS-14 variant peptide"  
XX  
PN CA2325169-A1.  
XX  
XX 03-JUN-2001.  
XX  
PF 01-DEC-2000; 2000CA-2325169.  
XX  
PR 03-DEC-1999; 99US-0168934.  
XX  
XX (NDSU-) NDSU RES FOUND.  
XX  
PI Sheridan MA, Moore CA, Kittelson JD;  
XX  
XX WPI; 2001-425997/46.  
DR N-PSDB; AAS12935.  
XX  
XX New somatostatin polypeptides derived from Oncorhynchus mykiss, useful  
PT for treating diabetes mellitus, acromegaly, gastrinoma, acquired  
PT immunodeficiency syndrome and neurological disorders.  
XX  
XX Claim 1; Fig 3; 52pp; English.  
XX  
XX The invention relates to an Oncorhynchus mykiss somatostatin polypeptide  
CC containing a portion of preprosomatostatin I (PPSS-I) and/or a portion of  
CC preprosomatostatin II (PPSS-II). The protein sequences and their  
CC associated polynucleotides are useful for identifying modified  
CC somatostatin polypeptide which functions as a somatostatin agonist useful  
CC for research, therapeutics or diagnostics, including medical and  
CC veterinary applications. The wild-type somatostatin and its modified  
CC version are useful for treating hypersecretion from endocrine tumours in  
CC the pituitary (e.g. acromegaly) or gastroenteropancreatic tissues (e.g.  
CC gastrinoma, glucagonoma, carcinoid syndrome), to cause tumour shrinkage  
CC through their effects on cell proliferation and apoptosis and as adjuncts  
CC in the treatment of diabetes mellitus via inhibition of growth hormone  
CC and glucagon. In addition, dysfunctional somatostatin secretion is  
CC associated with acquired immunodeficiency syndrome (AIDS) and various  
CC neurological disorders (e.g. epilepsy, Alzheimer's disease and  
CC Huntington's disease) and somatostatin antagonists are effective in the  
CC treatment of such conditions. Nucleic acids encoding the polypeptides are  
CC useful in gene therapy and fusion peptides can be targeted to neoplasms  
CC and their metastases, inhibiting the release of their secretory products.  
CC This sequence represents O. Mykiss PPSS-II' protein.  
CC Note: The features for this sequence are specifically claimed in the  
XX specification.  
XX  
SQ Sequence 111 AA;  
Query Match 100.0%; Score 122; DB 22; Length 111;  
Best Local Similarity 100.0%; Pred. No. 1e-10;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 MRVSIHICALALLGLALICSGAA 25  
Db 1 MRVSIHICALALLGLALICSGAA 25  
RESULT 2  
AAU07667  
ID AAU07667 standard; Protein; 115 AA;  
XX  
AC AAU07667;  
XX  
DT 04-DEC-2001 (first entry)  
XX  
DE Rainbow trout preprosomatostatin II (PPSS-II') polypeptide.  
XX  
KW Rainbow trout; somatostatin; preprosomatostatin; hypersecretion; PPSS-I;  
KW PPSS-II'; PPSS-II'; endocrine tumour; pituitary gland; glucagonoma; AIDS;  
KW gastroenteropancreatic tissue; acromegaly; gastrinoma; diabetes mellitus;  
KW carcinoid syndrome; cell proliferation; apoptosis; growth hormone;

KW glucagon; acquired immunodeficiency syndrome; neurological disorder; HIV;  
KW epilepsy; Alzheimer's disease; Huntington's disease; neuroprotective;  
KW neoplasm; metastasis; gene therapy; antidiabetic; nootropic; cytostatic;  
KW anti-human immunodeficiency virus; osteopathic; anticonvulsant.  
XX  
OS Oncorhynchus mykiss.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..25 /note= "Signal peptide"  
FT Protein 1..87 /note= "PPSS-II' pre-sequence"  
FT Protein 26..115 /note= "Mature PPSS-II'"  
FT Misc-difference 74 /note= "Encoded by CAA"  
FT Peptide 88..101 /note= "PPSS-II' pro-sequence"  
FT Peptide 88..115 /note= "Prosomatostatin II'"  
FT Cleavage-site 100..101 /note= "Dibasic cleavage site"  
FT Peptide 102..115 /note= "SS-14 variant peptide"  
XX  
XX CA2325169-A1.  
XX  
XX 03-JUN-2001.  
XX  
XX 01-DEC-2000; 2000CA-2325169.  
XX  
XX 03-DEC-1999; 99US-0168934.  
XX (NDSU-) NDSU RES FOUND.  
XX Sheridan MA, Moore CA, Kittelson JD;  
XX WPI; 2001-425997/46.  
DR N-PSDB; AAS12934.  
XX  
XX New somatostatin polypeptides derived from Oncorhynchus mykiss, useful  
PT for treating diabetes mellitus, acromegaly, gastrinoma, acquired  
PT immunodeficiency syndrome and neurological disorders.  
XX  
XX Claim 2; Fig 3; 52pp; English.  
XX  
XX The invention relates to an Oncorhynchus mykiss somatostatin polypeptide  
CC containing a portion of preprosomatostatin I (PPSS-I) and/or a portion of  
CC preprosomatostatin II (PPSS-II). The protein sequences and their  
CC associated polynucleotides are useful for identifying modified  
CC somatostatin polypeptide which functions as a somatostatin agonist useful  
CC for research, therapeutics or diagnostics, including medical and  
CC veterinary applications. The wild-type somatostatin and its modified  
CC version are useful for treating hypersecretion from endocrine tumours in  
CC the pituitary (e.g. acromegaly) or gastroenteropancreatic tissues (e.g.  
CC gastrinoma, glucagonoma, carcinoid syndrome), to cause tumour shrinkage  
CC through their effects on cell proliferation and apoptosis and as adjuncts  
CC in the treatment of diabetes mellitus via inhibition of growth hormone  
CC and glucagon. In addition, dysfunctional somatostatin secretion is  
CC associated with acquired immunodeficiency syndrome (AIDS) and various  
CC neurological disorders (e.g. epilepsy, Alzheimer's disease and  
CC Huntington's disease) and somatostatin antagonists are effective in the  
CC treatment of such conditions. Nucleic acids encoding the polypeptides are  
CC useful in gene therapy and fusion peptides can be targeted to neoplasms  
CC and their metastases, inhibiting the release of their secretory products.  
CC This sequence represents O. Mykiss PPSS-II' protein.  
CC Note: The features for this sequence are specifically claimed in the  
XX specification.  
XX  
SQ Sequence 115 AA;  
Query Match 90.2%; Score 110; DB 22; Length 115;  
Best Local Similarity 88.0%; Pred. No. 6.7e-09;







Query Match 42.6%; Score 52; DB 18; Length 433;  
Best Local Similarity 58.8%; Pred. No. 12;  
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALAICSQGA 25  
| : : : : :  
Db 107 ANAILGVSLAVCKAGAA 123

RESULT 5  
AAM48922  
ID AAM48922 standard; Protein; 434 AA.

XX AC AAM48922;

XX DT 16-APR-2002 (first entry)

XX DE Human enolase 3(beta,muscle) protein.

XX KW Human; enolase 3(beta, muscle); ENO3; single nucleotide polymorphism;  
XX KW SNP; haplotype analysis; isogene.

XX OS Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 71 /note= "optionally Ser"

FT Misc-difference 76

FT Misc-difference 85 /note= "optionally Pro"

FT Misc-difference 85 /note= "optionally Val"

FT Misc-difference 332 /note= "optionally Ile"

XX WO200202579-A2.

XX PN 10-JAN-2002.

XX PD 02-JUL-2001; 2001WO-US20952.

XX PF 30-JUN-2000; 2000US-215236P.

XX PR (GENA-) GENAISSANCE PHARM INC.

XX PA Duda A, Finkel K, Koshy B, Parks KE;

XX PI WPI; 2002-154721/20.

XX DR N-PSDB; AAK98531, AAK98532.

XX PT Novel genetic variants of enolase 3, (beta, muscle) gene useful in  
studying expression and function of the protein, and for screening  
drugs to treat disorders of glycolytic pathway

PS Claim 28; Fig 3; 90pp; English.

CC The present invention provides the protein, cDNA and genomic sequences of  
a human enolase 3 (beta, muscle) isogene containing a number of single  
nucleotide polymorphisms (SNPs). The sequences can be used to identify  
the haplotype of an individual and identify whether particular haplotypes  
are linked to certain diseases. The present sequence is the ENO3  
protein.

XX Sequence 434 AA;

Query Match 42.6%; Score 52; DB 23; Length 434;  
Best Local Similarity 58.8%; Pred. No. 12;  
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALAICSQGA 25  
| : : : : :  
Db 108 ANAILGVSLAVCKAGAA 124

RESULT 6  
AAM81913  
ID AAM81913 standard; Protein; 120 AA.

XX AC AAM81913;

XX DT 13-NOV-2001 (first entry)

XX DE Human haematological malignancy-related antigen #1611.

XX KW Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;  
XX KW haematological malignancy; antigen; chronic lymphocytic leukaemia;  
XX KW follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma.

XX OS Homo sapiens.

XX PN WO200164886-A2.

XX PD 07-SEP-2001.

XX PF 01-MAR-2001; 2001WO-US07272.

XX PR 01-MAR-2000; 2000US-0186126.

XX PR 17-MAR-2000; 2000US-0190479.

XX PR 27-APR-2000; 2000US-0200545.

XX PR 28-APR-2000; 2000US-0200303.

XX PR 01-MAY-2000; 2000US-0200779.

XX PR 04-MAY-2000; 2000US-0200999.

XX PR 22-MAY-2000; 2000US-0206201.

XX PR 14-JUL-2000; 2000US-0218950.

XX PR 03-AUG-2000; 2000US-0222903.

XX PR 04-AUG-2000; 2000US-0223416.

XX PR 07-AUG-2000; 2000US-0223378.

XX PA (CORI-) CORIXA CORP.

XX PI Gaiger A, Algate PA, Mannion J;

XX WPI; 2001-514842/56.

XX PT Compositions and methods for the detection of hematological  
malignancies, e.g. chronic lymphocytic leukemia, lymphoma, follicular  
lymphoma and Hodgkin's and T/B'cell non-Hodgkin's lymphoma

XX PS Claim 1; Page 1152; 1252pp; English.

XX CC The present invention relates to compositions and methods for the  
detection, diagnosis and therapy of haematological malignancies. The  
present sequence is the protein sequence of a human haematological  
malignancy related antigen. The methods of the present invention comprise  
detecting the presence of haematological malignancy related antigen(s) in  
a sample obtained from the patient (an increased level of the  
polypeptide, compared to an unaffected individual, is indicative of an  
increased risk). Haematological malignancies which can be treated using  
the present invention are chronic lymphocytic leukaemia, lymphoma,  
follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B  
cell non-Hodgkin's lymphoma.

XX Sequence 120 AA;

Query Match 39.3%; Score 48; DB 22; Length 120;  
Best Local Similarity 56.2%; Pred. No. 12;  
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALAICSQGA 24

Db 65 ANAILGVSLAVCKAGA 80

RESULT 7

AAM81978

ID AAM81978 standard; Protein; 162 AA.







QY 9 ALALLGLALAICSQGA 24  
Db 103 ANAILGVSLAVCKAGA 118  
RESULT 9  
AAW54357  
ID AAW54357 standard; protein; 433 AA.  
XX  
AC AAW54357;  
XX  
DT 14-AUG-1998 (first entry)  
XX  
DE Alpha Enolase.  
XX  
KW Endometrium; hyperplasia; adenocarcinoma; proliferative phase;  
KW 2D gel electrophoresis; detection.  
XX  
OS Homo sapiens.  
XX  
PN WO9810291-A1.  
PD 12-MAR-1998.  
XX  
PR 05-SEP-1997; 97WO-GB02394.  
PR 08-APR-1997; 97GB-0007132.  
PR 06-SEP-1996; 96GB-0018600.  
XX  
PA (CLIN-) CENT CLINICAL & BASIC RES.  
XX  
PI Byrjalsen I, Fey SJ, Larsen P;  
XX  
DR WPI; 1998-207057/18.  
XX  
PT Biochemical markers of human endometrium - useful for, e.g.  
PT diagnosis of hyperplasia and adenocarcinoma  
XX  
PS Disclosure; Page 21; 77pp; English.  
XX  
CC Proteins AAW54349-W54364 are examples of proteins produced in the  
CC endometrium during the hyperplasia, adenocarcinoma or proliferative  
CC phase of the endometrium. The presence and quantities of these proteins  
CC can be detected using 2D gel electrophoresis comparison of cell lysates.  
CC The proteins can be used as biochemical markers to detect the phase of  
CC the endometrium and can be measured in body fluids, obviating the need  
CC for endometrial biopsies.  
XX  
SQ Sequence 433 AA;  
Query Match 39.3%; Score 48; DB 19; Length 433;  
Best Local Similarity 56.2%; Pred. No. 48;  
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
QY 9 ALALLGLALAICSQGA 24  
Db 107 ANAILGVSLAVCKAGA 122  
RESULT 10  
ABB57379  
ID ABB57379 standard; Protein; 434 AA.  
XX  
AC ABB57379;  
XX  
DT 08-MAR-2002 (first entry)  
XX  
DE Rat mucocardial cell proliferation associated polypeptide SEQ ID NO 10.  
XX  
KW Rat; heart; cardiant; myocardial necrosis; cardiac hypertrophy;  
KW cardiac insufficiency.  
XX  
OS Rattus norvegicus.

XX  
PN WO200183705-A1.  
XX  
PD 08-NOV-2001.  
XX  
PF 27-APR-2001; 2001WO-JP03700.  
XX  
PR 27-APR-2000; 2000JP-0126741.  
XX  
PA (KYOW ) KYOWA HAKKO KOGYO KK.  
XX  
PI Yamada Y, Sekine S, Kikuchi Y, Sakurada K;  
XX  
DR WPI; 2002-075160/10.  
DR N-PSDB; ABI99919.  
XX  
PT Genes having differential expression in fetal and adult heart tissue  
PT useful for screening potential drugs for promoting repair of damage  
PT caused by myocardial necrosis  
XX  
PS Claim 53; Page 98-100; 171pp; Japanese.  
XX  
CC The invention relates to gene sequences (ABI99915-ABI99934) having  
CC modified expression in fetal heart tissue as compared to adult heart  
CC tissue and the encoded proteins (ABB57375-ABB57392). The genes have  
CC cardiant activity and may be useful in the promotion of the repair of  
CC damage to heart tissue caused by myocardial necrosis. The gene sequences  
CC are useful for screening potential compounds for the ability to influence  
CC disease associated with myocardial necrosis. Drugs identified by the  
CC screening methods may be used to treat and prevent disease with which  
CC myocardial necrosis is associated, such as cardiac hypertrophy and  
CC cardiac insufficiency. Diagnosis of diseases such as those above is also  
CC disclosed.  
XX  
SQ Sequence 434 AA;  
Query Match 39.3%; Score 48; DB 23; Length 434;  
Best Local Similarity 56.2%; Pred. No. 49;  
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
QY 9 ALALLGLALAICSQGA 24  
Db 108 ANAILGVSLAVCKAGA 123  
RESULT 11  
AAV35223  
ID AAV35223 standard; Protein; 210 AA.  
XX  
AC AAV35223;  
XX  
DT 13-SEP-1999 (first entry)  
XX  
DE Chlamydia pneumoniae protein not found in C. trachomatis.  
XX  
KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;  
KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;  
KW vaccine; neutralising epitope.  
XX  
OS Chlamydia pneumoniae.  
XX  
PN WO9927105-A2.  
XX  
PD 03-JUN-1999.  
XX  
PF 20-NOV-1998; 98WO-IB01890.  
XX  
PR 04-NOV-1998; 98US-0107078.  
PR 21-NOV-1997; 97FR-0014673.  
XX  
PA (GEST ) GENSET.  
XX  
PI Griffiths R;



XX WPI; 1999-357842/30.  
XX Genome sequence of Chlamydia pneumoniae.  
XX Page 1070-1071; Disclosure; 1912pp; English.  
XX AAY34584-Y35879 represent the proteins encoded by all the open reading  
CC frames in the complete genome (see AAX91990) of Chlamydia pneumoniae.  
CC C. pneumoniae causes respiratory disease such as pneumonia and  
CC bronchitis and is thought to be a contributing factor in heart  
CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema  
CC nodosum or pharyngitis. The polypeptides encoded by the open reading  
CC frames of the C. pneumoniae genome (see AAY34584-Y35879) can be used in  
CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae  
CC nucleotide sequences can also be used as immunogenic compositions,  
CC especially where the vector directs the expression of a neutralising  
C epitope of C. pneumoniae.  
XX Sequence 210 AA;  
SQ  
Query Match 38.5%; Score 47; DB 20; Length 210;  
Best Local Similarity 45.0%; Pred. No. 32;  
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;  
QY 5 QIHCALALLGLALICSGA 24  
Db 48 EIASAAILGLLVAFCSAA 67  
RESULT 12  
AAB47765  
ID AAB47765 standard; Protein; 423 AA.  
XX AAB47765;  
XX 21-FEB-2002 (first entry)  
DT Novel G-protein coupled receptor #5.  
DE Novel G-protein coupled receptor; NGPCR; transmembrane domain;  
XX signal transduction; ligand binding; human; placenta; bone marrow;  
KW trachea; testis; liver; kidney cell.  
KW Homo sapiens.  
XX WO200181411-A2.  
XX 01-NOV-2001.  
XX 27-APR-2001; 2001WO-US13371.  
XX 27-APR-2000; 2000US-199950P.  
XX (LEXI-) LEXICON GENETICS INC.  
PI Wang X, Walke DW, Scoville J;  
XX WPI; 2002-049266/06.  
DR N-PSDB; AAI72010.  
XX Novel isolated G-protein coupled receptor polypeptide useful for  
PT identifying compounds that modulate the protein expression or activity  
PT which are used to treat a variety of physiological or mental disorders  
PT  
XX Disclosure; Page 78-79; 85pp; English.  
XX The sequences given in AAB47761-69 show novel G-protein coupled  
CC receptor (NGPCR) polypeptides. The NGPCR coding sequences are found on  
CC chromosome 3 at 3q11.1 in the human genome. These NGPCR proteins show  
CC the presence of 7 conserved transmembrane domains which are  
CC interconnected by non-conserved hydrophilic loops. They are involved

CC in signal transduction after ligand binding. These NGPCR's are  
CC expressed in human placenta, bone marrow, trachea, testis, liver and  
XX kidney cells.  
SQ Sequence 423 AA;  
Query Match 38.5%; Score 47; DB 23; Length 423;  
Best Local Similarity 45.0%; Pred. No. 67;  
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;  
QY 3 VSQIHCALALLGLALICSQ 22  
Db 142 LSNVGCALSVTGLALTIVFQ 161  
RESULT 13  
AAU93172  
ID AAU93172 standard; Protein; 472 AA.  
XX AAU93172;  
XX 02-JUL-2002 (first entry)  
DT Arabidopsis transcription factor #210.  
DE Arabidopsis transcription factor #210.  
XX  
KW Agriculture; metabolic chemical; environmental stress; drought;  
KW microbial disease resistance; herbicide resistance; seed yield;  
KW fruit yield; growth rate; leaf senescence; flower senescence.  
KW plant; transcription factor; transgenic.  
XX Arabidopsis thaliana.  
OS Arabidopsis thaliana.  
XX WO200215675-A1.  
XX 28-FEB-2002.  
XX 22-AUG-2001; 2001WO-US26189.  
XX 22-AUG-2000; 2000US-227439P.  
PR 16-NOV-2000; 2000US-0713994.  
PR 16-APR-2001; 2001US-0837944.  
XX (MEND-) MENDEL BIOTECHNOLOGY INC.  
PA (PILG/) PILGRIM M.  
PA (CREE/) CREELMAN R.  
PA (DUBE/) DUBELL A J.  
PA (HEAR/) HEARD J.  
PA (JIAN/) JIANG C.  
PA (KEDD/) KEDDIE J.  
PA (ADAM/) ADAM L.  
PA (RATC/) RATCLIFF O.  
PA (REUB/) REUBER J L.  
PA (RIEC/) RIECHMANN J L.  
PA (YUGG/) YU G.  
PA (PINE/) PINEDA O.  
XX Pilgrim M, Creelman R, Dubell AJ, Heard J, Jiang C, Keddle J;  
PI Adam L, Ratcliff O, Reuber JL, Riechmann JL, Yu G, Pineda O;  
XX WPI; 2002-292022/33.  
DR N-PSDB; ABK65358.  
XX An isolated or recombinant polynucleotide used to produce a transgenic  
PT plant  
PT  
XX Claim 40; Page 864-866; 941pp; English.  
XX The invention relates to 1 of 232 isolated or recombinant polynucleotides  
CC encoding an Arabidopsis thaliana transcription factor, their variants,  
CC complements, fragments, or related polynucleotide with 31% to 95%  
CC sequence identity, where the plant possesses an altered trait as compared  
CC to a wild-type or reference plant, or the plant exhibits an altered  
CC phenotype as compared to a wild-type or reference plant, or the plant



CC exhibits ectopic expression or altered expression of one or more genes  
CC associated with a plant trait as compared to a wild plant. Also included  
CC are a transgenic plant comprising the polynucleotides, a computer  
CC readable medium having stored sequence information, and identifying a  
CC homologue sequence from a database comprising a plurality of known plant  
CC sequences comprising inputting sequence information selected from one of  
CC 464 fully defined sequences given in the specification. The isolated or  
CC recombinant polynucleotide is used for producing a plant having a  
CC modified trait, the method comprising selecting a polynucleotide that  
CC encodes a polypeptide or an antisense nucleic acid, inserting the  
CC polynucleotide or antisense nucleic acid into an expression vector,  
CC introducing the vector into a plant or a cell of a plant to overexpress  
CC the polypeptide or antisense nucleic acid, thereby producing a modified  
CC plant, and selecting for a modified trait (e.g. increased  
CC production of agriculturally useful proteins or metabolic chemicals,  
CC pest tolerance, environmental stress response (e.g. drought), microbial  
CC disease resistance, herbicide resistance, seed and fruit yield, growth  
CC rate, leaf and flower senescence and many other traits listed in the  
CC specification). The present sequence is one of the 232 proteins which are  
CC A. thaliana transcription factors.

XX SQ Sequence 472 AA;

Query Match 38.5%; Score 47; DB 23; Length 472;  
Best Local Similarity 40.9%; Pred. No. 75;  
Matches 9; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 4 SQIHCALALLGLALAICSGAA 25  
Db 193 AQIHAASVAGVAAVAIAAA 214

RESULT 14

AAU93018  
ID AAU93018 standard; Protein; 475 AA.  
XX AAU93018;

DT 02-JUL-2002 (first entry)

DE Arabidopsis transcription factor #56.

XX Agriculture; metabolic chemical; environmental stress; drought;  
KW microbial disease resistance; herbicide resistance; seed yield;  
KW fruit yield; growth rate; leaf senescence; flower senescence.  
KW plant; transcription factor; transgenic.

XX Arabidopsis thaliana.

OS WO200215675-A1.

PN 28-FEB-2002.

XX 22-AUG-2001; 2001WO-US26189.

PR 22-AUG-2000; 2000US-227439P.

PR 16-NOV-2000; 2000US-0713994.

PR 16-APR-2001; 2001US-0837944.

XX (MEND-) MENDEL BIOTECHNOLOGY INC.

PA (PILG/) PILGRIM M.

PA (CREE/) CREELMAN R.

PA (DUBE/) DUBELL A J.

PA (HEAR/) HEARD J.

PA (JIAN/) JIANG C.

PA (KEDD/) KEDDIE J.

PA (ADAM/) ADAM L.

PA (RATC/) RATCLIFF O.

PA (REUB/) REUBER J L.

PA (RIEC/) RIECHMANN J L.

PA (YUGG/) YU G.

PA (PINE/) PINEDA O.

XX

PI Pilgrim M, Creelman R, Dubell AJ, Heard J, Jiang C, Keddie J;  
PI Adam L, Ratcliff O, Reuber JL, Riechmann JL, Yu G, Pineda O;  
XX WPI; 2002-292022/33.  
DR N-PSDB; ABK65204.

XX An isolated or recombinant polynucleotide used to produce a transgenic plant

XX Claim 40; Page 265-267; 94lpp; English.

XX The invention relates to 1 of 232 isolated or recombinant polynucleotides encoding an Arabidopsis thaliana transcription factor, their variants, complements, fragments, or related polynucleotide with 31% to 95% sequence identity, where the plant possesses an altered trait as compared to a wild-type or reference plant, or the plant exhibits an altered phenotype as compared to a wild-type or reference plant, or the plant exhibits ectopic expression or altered expression of one or more genes associated with a plant trait as compared to a wild plant. Also included are a transgenic plant comprising the polynucleotides, a computer readable medium having stored sequence information, and identifying a homologue sequence from a database comprising a plurality of known plant sequences comprising inputting sequence information selected from one of 464 fully defined sequences given in the specification. The isolated or recombinant polynucleotide is used for producing a plant having a modified trait, the method comprising selecting a polynucleotide that encodes a polypeptide or an antisense nucleic acid, inserting the polynucleotide or antisense nucleic acid into an expression vector, introducing the vector into a plant or a cell of a plant to overexpress the polypeptide or antisense nucleic acid, thereby producing a modified plant, and selecting for a modified trait (e.g. increased production of agriculturally useful proteins or metabolic chemicals, pest tolerance, environmental stress response (e.g. drought), microbial disease resistance, herbicide resistance, seed and fruit yield, growth rate, leaf and flower senescence and many other traits listed in the specification). The present sequence is one of the 232 proteins which are A. thaliana transcription factors.

XX SQ Sequence 475 AA;

Query Match 38.5%; Score 47; DB 23; Length 475;

Best Local Similarity 40.9%; Pred. No. 75;

Matches 9; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 4 SQIHCALALLGLALAICSGAA 25

Db 198 AQIHAASVAGVAAVAIAAA 219

RESULT 15

AAU93018

ID AAU93018 standard; Protein; 481 AA.

XX AAU93018;

DT 21-FEB-2002 (first entry)

DE Novel G-protein coupled receptor #4.

XX Novel G-protein coupled receptor; NGPCR; transmembrane domain;

XX signal transduction; ligand binding; human; placenta; bone marrow;

XX trachea; testis; liver; kidney cell.

XX Homo sapiens.

XX WO200181411-A2.

PN 01-NOV-2001.

PD 27-APR-2001; 2001WO-US13371.

PR 27-APR-2000; 2000US-199950P.

XX



PA (LEXI-) LEXICON GENETICS INC.  
XX  
PI Wang X, Walke DW, Scoville J;  
XX  
DR WPI; 2002-049266/06.  
DR N-PSDB; AAI72009.  
XX  
PT Novel isolated G-protein coupled receptor polypeptide useful for  
PT identifying compounds that modulate the protein expression or activity  
PT which are used to treat a variety of physiological or mental disorders  
PT  
XX  
PS Claim 6; Page 77; 85pp; English.  
XX  
XX The sequences given in AAB47761-69 show novel G-protein coupled  
CC receptor (NGPCR) polypeptides. The NGPCR coding sequences are found on  
CC chromosome 3 at 3q11.1 in the human genome. These NGPCR proteins show  
CC the presence of 7 conserved transmembrane domains which are  
CC interconnected by non-conserved hydrophilic loops. They are involved  
CC in signal transduction after ligand binding. These NGPCR's are  
CC expressed in human placenta, bone marrow, trachea, testis, liver and  
CC kidney cells.  
XX  
SQ Sequence 481 AA;

Query Match 38.5%; Score 47; DB 23; Length 481;  
Best Local Similarity 45.0%; Pred. No. 76;  
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 3 VSQIHGALALLGLAICSQ 22  
:|:|::|::|  
Db 200 LSNVGCALSVTGLATVIFQ 219

Search completed: March 21, 2003, 11:38:04  
Job time : 17.4506 secs